

It's not all about grades: new perspectives on graduate students' academic performance

Citation for published version (APA):

Schwager, I. T. L. (2015). It's not all about grades: new perspectives on graduate students' academic performance. [Doctoral Thesis, Maastricht University]. Maastricht University.
<https://doi.org/10.26481/dis.20151125is>

Document status and date:

Published: 01/01/2015

DOI:

[10.26481/dis.20151125is](https://doi.org/10.26481/dis.20151125is)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

***It's Not All About Grades:
New Perspectives on Graduate Students' Academic Performance***

by

Inge Thekla Louise Schwager

COLOPHON

The research presented in this dissertation was supported by a grant of Educational Testing Service (ETS), Princeton, New Jersey. ETS provided access to the ETS PPI and the GRE® revised General Test. ETS and GRE are registered trademarks of Educational Testing Service (ETS) in the United States and other countries.

Cover design: Alexander Vaccaro

© 2015 Inge Thekla Louise Schwager, Köln

All rights reserved. No parts of this thesis may be reproduced or transmitted in and form or by any means, electronic or mechanical, including photocopy, recording or any other information storage or retrieval system, without permission in writing from the author, or, when appropriate, from the publishers holding the copyright of the publishes articles.

It's Not All About Grades:

New Perspectives on Graduate Students' Academic Performance

DISSERTATION

to obtain the degree of Doctor at Maastricht University, on the authority of the Rector Magnificus,
Prof. Dr. L.L.G. Soete in accordance with the decision of the Board of Deans, to be defended in public
on Wednesday 25th of November 2015 at 12.00 hours by

by

Inge Thekla Louise Schwager

Supervisor:

Prof. Dr. F.R.H. Zijlstra

Co-supervisor:

Dr. U.R. Hülshager

Assessment Committee:

Prof. Dr. G. Kok (chairman)

Prof. Dr. M.Ph. Born (Erasmus University Rotterdam)

Prof. Dr. C.J. König (Saarland University)

Prof. Dr. H.L.G.J. Merckelbach

Prof. Dr. R.K.W. van der Velden

Table of Contents

Chapter 1	General Introduction	8
Chapter 2	Graduate Student Selection: GRE, Socioeconomic Status and Undergraduate Grade Point Average as Predictors of Study Success in a Western European University	34
Chapter 3	Supervisor Ratings of Students' Academic Potential as Predictors of Citizenship and Counterproductive Behavior	54
Chapter 4	Be Aware to Be on the Square: Mindfulness and Counterproductive Academic Behavior	73
Chapter 5	General Discussion	89
	References	102
	Appendix	119
	Valorization	124
	Summary	127
	Samenvatting	133
	Acknowledgements	136
	Curriculum Vitae	139

Chapter 1

General Introduction

In 1999, the Ministers of Education of 29 European countries signed the Bologna declaration and started what has been known as the Bologna process (Education, Audiovisual and Culture Executive Agency, 2010). In response to the Bologna protocol, European universities introduced a comprehensive educational system with a bachelor and master phase. One cornerstone of the Bologna process was to be more responsive to the needs of the labor market (Terry, 2008). In doing so, universities have been encouraged to strengthen students' employability which can be defined as "a set of achievements – skills, understandings and personal attributes – that makes graduates more likely to gain employment and be successful in their chosen occupations, which benefits themselves, the workforce, the community and the economy" (Yorke, 2006; p. 8). Going beyond knowledge transfer by contributing to students' sustainable employability clearly broadens the scope of universities. However, they simultaneously increase the likelihood of educating a student body with a positive influence on their own well-functioning and reputation.

The idea that students' employability depends on additional factors than solely acquiring academic knowledge is in line with recent research conducted in the industrial and organizational (I/O) psychology literature defining performance as a multidimensional concept encompassing not only task performance, but also citizenship and counterproductive behavior (Borman & Motowildo, 1993; Rotundo & Sackett, 2002). Whereas, task performance is defined as "activities that are formally recognized as part of the job and that contribute to the organization's technical core" (Borman & Motowildo, 1993, p. 73), citizenship performance is "the maintenance and enhancement of the social and psychological context that supports task performance" (Organ, 1997, p. 91). Counterproductive behavior, in contrast, is defined as "behavior that harms the well-being of the organization" (Rotundo & Sackett, 2002, p. 69). Citizenship and counterproductive behaviors are both considered to be non-task related performance aspects. Previous research provides evidence for the distinctiveness of the three performance dimensions (Dalal, 2005; Hoffman, Blair, Meriac, & Woehr, 2007; Spector, Bauer & Fox, 2010). Thus, universities committed to the aim of achieving sustainable employability in their student body are encouraged to use admission criteria covering all three performance aspects. Furthermore, they should think about appropriate means to further develop students once they have been admitted. The overall goal of the present dissertation is therefore twofold: First, I aim to assess the validity of two selection methods in predicting student task performance, citizenship and counterproductive behavior. Second, I will zoom in on counterproductive behavior and investigate the role of a number of selected personality traits in counterproductive academic behavior.

Selecting Students with High Task and Non-Task Performance

Although past behavior is a valid predictor of future behavior (Wernimont & Campbell, 1968), it is not advisable to rely exclusively on undergraduate grade point average (U-GPA) when predicting students' task performance measured in terms of graduate grade point average (G-GPA). Students' mobility has considerably increased between European countries. In the Netherlands, for instance, the amount of foreign students has grown by 28 percent (45 thousand students from abroad) between 2005 and 2009 (Statistics Netherlands, 2010). Thus, applying students come from a broad range of countries using markedly different grading systems and differing in the leniency and strictness of grades. As a consequence, U-GPAs are hardly comparable to each other, lack objectivity and may thereby lead to unfair selection decisions. An obvious solution for European universities to select well-suited master students would be to adopt internationally established standardized tests assessing cognitive skills that predict future academic success. A major advantage of standardized tests is that they allow for a direct comparison of knowledge and academic potential of students with different educational backgrounds. Over the last four decades, a solid amount of literature has accumulated, documenting the validity of standardized tests as predictors of graduate school performance (Kuncel & Hezlett, 2007; Kuncel, Hezlett, & Ones, 2001). Most of this research has been conducted in the United States, where the majority of institutions use standardized tests to select their graduate students. European institutions, in contrast, have rarely been using information from standardized tests to make selection decisions at the graduate level. Consequently, evidence on the validity of standardized tests for selection of graduate students is still scarce in Europe. Validity evidence obtained in the United States, however, should not be generalized to other cultures and countries without further testing, as cultural differences and differences in educational systems may affect validities of selection methods (Hülshager, Maier, & Stumpp, 2007; Salgado & Anderson, 2002). It is therefore necessary to investigate the validity of standardized tests in the context of Europe, in order to assess the utility of such methods for graduate student selection at European universities. Therefore, the first aim of this dissertation is to investigate the predictive validity of a well-established cognitive test in order to predict students' task performance in the European context.

As previously mentioned educational and subsequent work performance manifests itself not only in grades, but in various additional outcomes such as social skills, perseverance and ethical behavior (Oswald, Schmitt, Kim, Ramsay, & Gillespie, 2004). Thus, a large amount of variance is likely left unexplained when one focuses solely on cognitive predictors (Schmitt et al., 2009).

Furthermore, qualities required in the educational context become more alike to qualities needed in the work context if students progress in their educational careers (Poropat, 2009). Thus, research conducted in the work context providing evidence for the utility of noncognitive predictors (Barrick & Mount, 1991), underlines the need to establish noncognitive assessments especially at a later educational stage. In response to that, researchers as well as practitioners have witnessed a current surge in interest in the use of supplemental noncognitive predictors in educational selection (e.g., Kyllonen, 2008; Kyllonen, Walters, & Kaufman, 2005). Whereas cognitively loaded tests yield the highest validities for task performance, predictors targeting noncognitive skills might yield the highest validities for contextual performance aspects (Côté & Miners, 2006; McHenry, Hough, Toquam, Hanson, & Ashworth, 1990; Motowildo, Borman, & Schmit, 1997; Schmitt et al., 2009). Thus, the second goal of the dissertation is to investigate the predictive validity of a noncognitive assessment tool in order to predict university citizenship and counterproductive academic behavior.

Identifying Mechanisms to Reduce Counterproductive Behavior

The detrimental impact of students' engaging in counterproductive behavior is twofold from universities' perspective. First, counterproductive academic behavior has a negative short-term influence on the daily well-functioning of the university and causes considerable costs. Second, the long-term employability of students with a high counterproductive behavior propensity is at risk. With respect to the negative short-term consequences, the biggest scandal has taken place at Harvard University in the last decade (Pérez-Peña, 2013). Substantial public interest resulted from this scandal due to which dozens of students had to temporarily withdraw from Harvard as they cheated on a take home exam. Another critical incident has taken place at Columbia University. Students stole Nutella from the university cafeteria in the amount of \$5.000 per week (Barron, 2013). Although this may sound funny at first, the costs associated with this behavior are immense from Columbia University's perspective. In the work context, the negative consequences of counterproductivity are yet more severe. Thus, employers aim to reduce the share of employees showing counterproductive behavior such as excessive absenteeism or theft. In doing so, they try to integrate predictors of counterproductive behavior into their selection procedures in order to screen out potential wrongdoers beforehand. Thus, universities caring about the future occupational success and long-term employability of their Alumni students are encouraged to underline and foster an awareness of integrity and honesty while they are still in charge of the students.

Academic dishonesty often goes along with other forms of deviant behavior such as risky driving, stealing, or cheating on income taxes (Blankenship & Whitley, 2000). Thus, a consistent pattern seems to underlie the tendency to engage in counterproductive behavior (Roberts, Harms, Caspi, & Moffitt, 2007). These findings suggest that counterproductive behavior is not only a function of the situation individuals find themselves in but also a function of the person. This is in line with empirical findings showing that stable personality traits such as conscientiousness and honesty-humility provide important insights into the occurrence of counterproductive academic behavior (de Vries, de Vries, & Born, 2011; Marcus, Lee & Ashton, 2007). However, personality traits as defined in the traditional personality taxonomies are relatively stable over time and cannot be readily influenced by external factors. Therefore, it is crucial to identify additional personality characteristics that can be altered by means of interventions. Mindfulness is such an alternative personality characteristic described as the capacity to be nonjudgmentally aware of the present moment (Brown & Ryan, 2003). Compared to less mindful students that are overwhelmed by their negative emotions due to stressful events in their everyday student life (e.g., an upcoming deadline), the present-moment orientation of individuals scoring high on mindfulness together with their nonjudgmental attitude might help them experiencing a stressful situation less threatening. This more objective approach might in turn evoke adequate response behaviors instead of maladaptive, counterproductive reactions. Therefore, mindfulness might contribute to the explanation of counterproductive behaviors while having the potential to be altered through interventions (Carmody & Baer, 2008; Shapiro, Brown, & Biegel, 2007). Thus, the third goal of this dissertation is to go beyond traditional personality predictors of counterproductive behavior and investigate the relationship between trait mindfulness and counterproductive behavior in the academic context.

Summing up, the overall aim of this dissertation is threefold (see Figure 1.1): First, I investigate the validity of a standardized cognitive admission test, namely the Graduate Record Examination (Chapter 2), in order to predict students' task performance. Second, I investigate the validity of a recently developed noncognitive assessment tool, the Personal Potential Index (ETS® PPI; Chapter 3), in order to predict university citizenship and counterproductive academic behavior. Third, I establish the relationship between trait mindfulness and counterproductive academic behavior (Chapter 4), as this might be a potential mechanism through which the occurrence of counterproductive behavior could be reduced. Whereas Chapter 2 and 3 have a strong focus on student selection by providing promising assessment tools that can be used to predict a broadened academic performance domain, Chapter 4 focuses on the role of personality traits, especially trait mindfulness for counterproductive academic behavior. It thereby paves the way for future research testing the effectiveness of mindfulness interventions in the academic context.

Taken together, this dissertation aims to contribute to a current trend in the literature to carefully match the predictor and the criterion domain (Borman, Brantley, & Hanson, 2014; Lievens, Buyse, & Sackett, 2005; Schmitt et al., 2009).

An extensive meta-analysis has recently shown that general mental ability (GMA) is a strong predictor of task performance, whereas personality traits are more appropriate to predict citizenship and counterproductive behavior (Gonzalez-Mulé, Mount, & Oh, 2014). By broadening and matching the predictor and the criterion domain, I aim to increase the fidelity of the predictive validity evidence obtained in this dissertation. Especially the usefulness of noncognitive predictors has been often underestimated in previous research when the predictive value has been only tested with respect to task performance (Ridgell & Lounsbury, 2004; Poropat, 2009; Richardson, Abraham, & Bond, 2012; Trapmann, Hell, Hirn, & Schuler, 2007). Thus, by underlining the importance of higher education as a precursor of future employment, this dissertation builds on a holistic conceptualization of performance. Thereby, I aim to provide new insights on how the different aspects of graduate students' academic performance can be assessed and potentially improved in such a way that both students and universities benefit from it.

This introduction is structured as followed: First, the idea to operationalize study performance in broader terms is elaborated. Second, the need to provide validity evidence for standardized tests in the European context is explained in more detail. Third, an overview about the current state of research on noncognitive predictors of study success is provided. Fourth, the relation between trait mindfulness and study related outcomes is established. Finally, the introduction is closed with an outline of this dissertation including a short description of each chapter.

Academic Performance - A Multidimensional Construct

Most people judge graduate students' study success in terms of grades or whether students obtain their master degree in the predefined amount of time. These are obviously two important aspects of academic performance from both students' and universities' perspective. Students' grades are especially crucial when they apply for a job in order to provide a proof of their achievement. From universities' perspective, students' grades are essential as they need an objective cut-off criterion based on which they provide students with a degree. The time students need to finish their master program is of special interest due to money related reasons. Universities calculate spending a certain amount of resources on each student. Staff members, for example, are considered to supervise a fixed number of master theses every academic year. In case many students do not finish their master thesis in time the number of students accumulates, tying up staff members' capacities. Thus, in order to warrant both the well-functioning of the university and students' graduation endeavors it seems straightforward that certain standards need to be defined in the curriculum based on which students take the decision whether they want to strive for a master degree.

However, students' performance encompasses more than just grades and degree attainment. Students work together in groups in which they share knowledge, learn from each other, have to agree upon solutions to solve complex problems, interact with staff members and so forth. They do not function independently from each other. In doing so, they can show behaviors that have a positive influence on the group functioning (e.g., do voluntary extra work), but they can also harm the group functioning by engaging in detrimental behavior (e.g., hold back material from peer students in order to be better informed). Furthermore, in interaction with university staff members' students considerably differ. Whereas some students use any given opportunity to criticize the circumstances at their university, others are much more lenient in their judgment and behavior. Assuming that staff members usually do not make mistakes on purpose (e.g., publish exam results with a delay) and that unpopular decisions sometimes need to be taken (e.g., a student is failing in an exam), the well-functioning of the university benefits from students that do not immediately engage in complaints, but take into account staff members' perspective as well. Dishonest behavior is also an issue in the academic context (McCabe, Butterfield, & Trevino, 2006). It can occur in various forms such as cheating on exams, plagiarism or fraudulent excuse making. Being aware of this problem, universities formulate ethical standards hoping to reduce the occurrence of dishonest behavior by increasing academic integrity. This is especially important as students' are more likely to show dishonest behavior when they perceive that other students engage in dishonest behavior as well (McCabe et al., 2006).

Thus, building on research in the work context (Borman & Motowildo, 1993; Rotundo & Sackett, 2002), researchers have started to acknowledge the role of citizenship and counterproductive behavior in the academic context. In one stream of literature, the authors focused on citizenship and counterproductive behavior as antecedents of study success. Allison, Voss and Dryer (2001) related university citizenship behavior to students' productivity and GPA, finding a significant positive relationship. Likewise, Credé and Niehorster (2009) could show that the occurrence of counterproductive academic behavior is negatively related with academic achievement in terms of grades. Recent work has focused on identifying the antecedents of citizenship and counterproductive behavior. Zettler (2011) established self-control as a valid predictor of university citizenship and counterproductive academic behavior. Furthermore, Meriac (2012) investigated the link between work ethic and citizenship and counterproductive behavior in the academic context. As hypothesized, work ethic was positively related with citizenship behavior and negatively related with cheating and disengagement.

Scholars in the I/O literature conceptualize citizenship and counterproductive behavior in different ways. For both citizenship and counterproductive behavior taxonomies exist differentiating between behaviors directed towards the organization and behaviors directed towards other individuals (Gruys & Sackett, 2003; Bennett & Robinson, 2000; Robinson & Bennett, 1995; Smith, Organ, & Near, 1983). However, we are not aware of empirical research conducted in the academic context that explicitly refers to this conceptualization of citizenship and counterproductive behavior. With respect to citizenship behavior, a taxonomy developed by Organ (1990) is more widely used. He divided the citizenship behavior construct into five dimensions specified as altruism, civic virtue, conscientiousness, courtesy, and sportsmanship. Allison and colleagues (2001) adapted the five dimensions to the academic context (see Table 1.1). Whereas sportsmanship, conscientiousness and civic virtue were valid predictors of students' productivity, sportsmanship and conscientiousness were predictive of students' GPA. Staufienbiel and Hartz (2000) could replicate only four of the five subdimensions in the European context as they did not find an empirical equivalent for the courtesy dimension. Thus, Zettler (2011) also omitted the courtesy dimension when he conducted research on citizenship behavior at a German university.

Table 1.1 Key Dimensions of Citizenship Behavior

Dimension	Description	Business setting	Academic setting example
Altruism	Voluntary actions that help a fellow employee with work-related problems.	Help fellow employees use equipment, complete work assignments, prepare for presentations, understand a computer software program, or locate information.	Help fellow students with their school work (e.g., complete an assignment, prepare for an examination, write a paper, run a computer program).
Civic Virtue	Voluntary participation in, and support of both a professional and social nature; in general, looking out for the organization's best interests.	Participate in organizational policy making and larger firm issues; attend optional meetings, forums, training sessions, etc.; monitor firm threats and opportunities; attend company-sponsored social events (e.g., company picnics); and assist with firm-supported philanthropic activities.	Volunteer to help organize or participate in school activities such as student government, campus social events, athletic team pep rallies, speakers series, philanthropic activities (e.g., food drives), and school club duties.
Conscientiousness	A pattern of going well beyond minimally required role and task requirements.	Arrive at work early and leave late; avoid prolonged or unnecessary breaks; be punctual for meetings and appointments; complete assignments before they are due; conserve company resources; make constructive suggestions; assume added responsibilities to help firm performance.	Attend class on time; turn assignments in early; participate in class discussion and activities with enthusiasm; and volunteer to do more work than is required, encouraging other students to do the same.
Courtesy	The discretionary enactment of thoughtful and considerate behaviors that prevent work-related problems for others.	Notify employer if one is going to be late or absent from work; notify co-workers in advance of committing to actions that will affect them; inform fellow employees of delays in work progress; or inform fellow workers of possible roadblocks in the completion of a project.	Inform instructors when unable to attend a class; inform team members when unable to attend meetings; obtain feedback from team members before making changes to team projects; or refrain from actions that would be disruptive to others during lectures.
Sportsmanship	A willingness to tolerate the inevitable inconveniences and impositions that result in an organization without complaining, and doing so with a positive attitude.	Refrain from complaining about having to work overtime to complete a project; having a deadline moved up; annoying, but not harmful, work conditions (e.g., uncomfortable temperature); or having one's ideas and suggestions rejected.	Refrain from complaining about instructor's delays in grading and assignment; classroom equipment malfunctions; or when class members do not contribute equally to team projects.

Retrieved from Allison, Voss, & Dryer (2001)

With respect to counterproductive academic behavior, Credé and Niehorster (2009) adhered to the idea of having two subfacets of counterproductive behavior (Robinson & Bennett, 1995) while making an important modification. They argue that students show counterproductive behavior directed towards others (e.g., property theft) and directed towards themselves (e.g., work avoidance). Using a questionnaire, designed to measure counterproductive tendencies in the academic context (Hakstian, Farrell, & Tweed, 2002), they found support for the two dimensional factor structure of counterproductive academic behavior. Constructs measuring individual differences such as self-control were more strongly related to self-focused counterproductive behavior than to other-focused counterproductive behavior.

Although citizenship and counterproductive behavior are considered to have different subfacets, recent studies investigating the construct validity of both constructs have shown that a single factor model fits the data well (Berry, Ones, & Sackett, 2007; Hoffman, et al., 2007; LePine, Erez, & Johnson, 2002; Poropat, 2009). In line with these findings justifying the creation of composite scores, a one-dimensional conceptualization of citizenship and counterproductive behavior is used in the remainder of this dissertation.

Summing up, in order to account for the multidimensional nature of performance, task performance is considered when validating the GRE, and a one-dimensional conceptualization of citizenship and counterproductive behavior when considering noncognitive predictors like the PPI and trait mindfulness.

The Need to Validate the GRE in the European Context

Especially in the US standardized tests are widely used in order to select those students that are most likely to succeed in educational settings. Based on the idea that students' cognitive abilities are a valid predictor of their study success several standardized tests have been developed. Whereas undergraduate students are selected based on Scholastic Aptitude Test (SAT) and American College Test (ACT) results, Graduate Management Admission Test (GMAT) and Graduate Record Examination (GRE) are used to select graduate students. Compared to the GMAT that was primarily developed to select business students, the GRE was designed to differentiate between students from all study fields. The GRE measures graduate students' intellectual capital by differentiating between three different aspects of ability namely verbal and quantitative reasoning and analytical writing.

Several meta-analyses (Goldberg & Alliger, 1992; Kuncel, et al., 2001; Kuncel, et al., 2009; Morrison & Morrison, 1995; Powers, 2004) have investigated the predictive validity of the GRE. Goldberg and Alliger (1992) conducted a meta-analysis integrating a comparatively small number of ten studies focusing on psychology students. The correlations between G-GPA and GRE was .15 for both verbal (GRE-V) and quantitative (GRE-Q) reasoning. Morrison and Morrison (1995) synthesized 22 studies across samples from various fields of study. They found higher correlations between G-GPA and GRE-V ($r = .28$) and GRE-Q ($r = .22$).

Meeting shortcomings of previous meta-analyses, an extensive meta-analysis was conducted including 1,521 studies (Kuncel, et al., 2001). In this study the authors corrected for range restriction and used several outcome measures. Furthermore, they conducted separate analyses for different study disciplines. The authors provided evidence for the validity of the three GRE parts in predicting G-GPA (GRE-V: $r = .34$; GRE-Q: $r = .32$; GRE-A: $r = .36$), faculty ratings (GRE-V: $r = .42$; GRE-Q: $r = .47$; GRE-A: $r = .35$), and comprehensive examination scores (GRE-V: $r = .44$; GRE-Q: $r = .26$; GRE-A: $r = .51$). The correlations with degree attainment (GRE-V: $r = .18$; GRE-Q: $r = .20$; GRE-A: $r = .11$) and time to degree attainment (GRE-V: $r = .28$; GRE-Q: $r = -.12$) were considerably smaller. This was probably due to noncognitive and situational variables which have a strong influence on these outcome measures.

In a second meta-analysis, Kuncel and colleagues (2009) differentiated between master and doctoral level in order to investigate the predictive validity of the GRE. Assuming that the predictive validity of the GRE might be altered by complexity, structure, selectivity and discipline of the study program, they hypothesized to find different correlations on master and doctoral level. Although being slightly different, GRE-V and GRE-Q predicted cumulative graduate grade point average at both, the master (GRE-V: $r = .38$; GRE-Q: $r = .30$), and the doctoral level (GRE-V: $r = .27$; GRE-Q: $r = .28$).

The meta-analyses cited above (Goldberg & Alliger, 1992; Kuncel, et al., 2001; Kuncel, et al., 2009; Morrison & Morrison, 1995) are based exclusively on primary studies conducted in North America, predominantly in the United States. In the field of personnel selection, a series of meta-analyses focusing on European primary studies have been conducted (Hülsheger et al., 2007; Salgado & Anderson, 2002; Salgado, Anderson, Moscoso, Bertua, & de Fruyt, 2003) to test whether the validity of general mental validity (GMA) for job performance and training success differs between North America and Europe.

Although all of these meta-analyses confirmed that GMA is a valid predictor, they also revealed validity differences between North America and Europe as well as between European countries. For instance, Salgado et al.'s (2003) meta-analysis of European studies revealed larger validities for the prediction of job performance than was previously reported in meta-analyses including primary studies conducted in North-America underscoring the need to conduct separate validity studies for the European context.

Thus, despite the aforementioned comprehensive validity evidence for the GRE, there is a need to investigate its validity in the European context as well. Given cultural and educational differences between North America and European countries, validities reported in these meta-analyses may not generalize to the European context. Educational systems, for example, are characterized by their degree of differentiation and tracking at the secondary school level. In stratified school systems students are sorted into different types of schools preparing them for different educational and occupational trajectories according to their former level of achievement (Van de Werfhorst, 2009). In those countries fewer students fulfill the requirements to enter postsecondary education. In less stratified educational systems comprehensive secondary schools are common and ability-based tracking is less widespread. Notably, stratification has an influence on the homogeneity of the student body. This in turn might influence the predictive validity of standardized tests explaining the need to conduct separate validity studies for countries that are characterized by different levels of stratification.

In the US where the majority of validity studies has been conducted, the educational system is characterized by a low level of stratification (Kerckhoff, 2001). High-schools enroll students independent of their individual abilities. Students graduating from secondary education then all have the same credentials, allowing equal access to postsecondary education. Thus, the student body entering postsecondary education is less homogenous compared to countries in which students are sorted into different types of education due to their achievement level at an early age.

The Dutch educational system is characterized by a stronger institutional differentiation (Müller & Shavit, 1998). Students can choose between three types of secondary education when they are approximately 12 years old. Pre-university education (VWO) takes six years and provides preparation for university, whereas general secondary education (HAVO) is designed to prepare students for higher professional education. Students who complete HAVO acquire ample qualification to proceed with VWO. Pre-vocational secondary education (VBO), however, offers basic vocational training and does not allow access to university.

The educational system with the strongest segregation between non-overlapping tracks is the German schooling system (Kerckhoff, 2001). Children are sorted into lower secondary school (Hauptschule), intermediate secondary school (Realschule), and upper secondary school (Gymnasium) when they are about 10 years old. Transition rates from elementary school to Gymnasium are approx. 30% (Schnabel, Alfeld, Eccles, Köller, & Baumert, 2002), and only those who successfully receive their University entrance certificate (Abitur) acquire formal permission to higher education. Those who finish Realschule can either continue in an advanced vocational school, or they choose to enter Gymnasium after completion of grade 10. The majority follows a three-year training program offered by companies and firms preparing them for a specific type of work. Students finishing Hauptschule are also enabled to do an apprenticeship. Although flexibility to switch between different school levels exists to a certain degree, it is more likely that students move from a higher to a lower level than vice versa (Heinz, 1999; Henz, 1997).

Taken together, European countries differ with respect to stratification at the secondary school level (Buchmann & Dalton, 2002; Müller & Shavit, 1998). Given that students originating from different European countries are mobile within Europe, it is difficult to predict whether the student body at a Dutch university is more or less homogenous compared to the US. However, separate validity evidence is needed to investigate whether potential differences in the student body alter GRE's validity in the European context.

Another important aspect that might influence the predictive validity of the GRE exists with respect to different grading traditions: European education systems in general and the Dutch system in particular differ from the US with respect to grading scheme and grading severity. While the US grading system is based on a 4-point scale ranging from 0.0 to 4.0, the Dutch grading system is based on a 10-point scale ranging from 1.0 to 10.0 with 10 representing perfect performance. Furthermore, US performance measures are frequently constricted because the majority of students receive 4s and 3s, especially on the graduate level (Johnson, 2003; Oldfield & Hutchinson, 1997). In contrast, the Netherlands has a strict grading culture in which 10s have hardly ever been awarded (Nuffic, 2012). Thus, the grade distribution is less skewed towards high grades. Both aspects may reduce range restriction and might result in higher validities in the Netherlands compared to the United States.

In addition, European and US applicant pools differ regarding their level of English proficiency. English is not the mother tongue for the majority of applicants to Dutch and other European master programs while the majority of applicants to US graduate schools are native English speakers. Previously, the validity of the GRE for non-native English speakers has been described as a concern (Kuncel et al., 2001) which might limit generalizability of validity evidence that has been gained in the United States. Although the study provided evidence that a similar pattern of results emerged for native and non-native speakers, researchers have called for investigations that take into account students' English proficiency as a factor that might influence the validity of the GRE. For validity studies conducted in the European context this is especially important given that most graduate students are non-native English speakers, except for students originating from UK.

Thus, in order to make sure that the GRE can be used as a predictor of study success in the European context, this dissertation sets out to investigate the predictive validity of the GRE at a Dutch university. In evaluating the GRE I focus on the prediction of study success operationalized in terms of grades. However, in order to cover contextual performance aspects, I focus on a newly developed noncognitive selection tool.

Noncognitive Predictors of Study Success

In the past, universities have predominantly relied on measures and indicators that target a student's cognitive potential, such as standardized tests (e.g., GRE or GMAT), or undergraduate grade point average (U-GPA). As mentioned above, these have been shown to be valid and useful in predicting students' task performance, as measured by future grade point average, faculty ratings, and comprehensive examination scores (Kuncel, et al., 2007; Kuncel, et al., 2001). However, noncognitive assessment tools explain additional variance when predicting students' performance that is not explained by cognitive predictors alone (Credé & Kuncel, 2008; Richardson, et al., 2012; Robbins et al., 2004). Furthermore, to predict an expanded criterion space including citizenship and counterproductive behavior, noncognitive predictors are considered to be more appropriate than cognitive predictors (McHenry, et al., 1990; Schmitt et al., 2003). An advantage of noncognitive compared to cognitive predictors is that they might reduce discrimination against minority groups. Whereas standardized tests such as the GRE still produce subgroup differences between women and men and between different ethnic groups, this discrimination does not necessarily translate into students' representation in higher education (Bleske-Rechek & Browne, 2014). This might be explained by the increased use of noncognitive assessments which allows students scoring lower on standardized tests counterbalancing their underperformance with better results in noncognitive assessments. Thus, in order to foster a diverse student body, stakeholders in admission testing are increasingly striving for noncognitive predictors that are time efficient while having sound psychometric properties.

Kyllonen (2008) has provided a comprehensive overview of different methods measuring noncognitive factors. Distinguishing between self-assessments, situational judgment tests (SJTs), biodata measures, objective measures and others' ratings, he discusses the advantages and disadvantages of the various methods. Although the classification itself is still up to date, new insights into the field of noncognitive assessments have emerged in previous years. Thus, in the following, I will complement this overview with recent research findings specifically focusing on self-assessments, SJTs, biodata measures and others' ratings (for an overview see Table 1.2).

The most prominent self-assessments predictive of academic performance are dispositional personality traits as operationalized in the Big Five and HEXACO taxonomy. Meta-analytical findings could show that conscientiousness, agreeableness, and openness are valid predictors of academic performance measured in terms of GPA (Poropat, 2009).

Notably, the predictive validity of conscientiousness was comparable to the predictive validity of intelligence. In order to predict citizenship behavior in the academic context, conscientiousness was also a valid predictor (Gore, Kiefner, & Combs, 2012). The same holds true for counterproductive academic behavior which is negatively related with conscientiousness as well as extraversion and honesty-humility (Marcus, et al., 2007). Furthermore, academic self-efficacy and achievement motivation (Robbins et al., 2004), grade goal (Richardson, et al., 2012) and attitudinal constructs such as study habits (Credé & Kuncel, 2008) explain additional variance above and beyond traditional antecedents of study success when predicting students GPA. Thus, selecting students by simply using a questionnaire assessing one of the aforementioned constructs might be promising on the first sight. However, as mentioned by Kyllonen (2008) self-assessments are prone to socially desirable responding. Thus, students might be able to respond in a social desirable manner in order to increase the likelihood of being admitted.

Situational judgment tests (SJT) are another possibility mentioned in the overview of Kyllonen (2008) having the potential to predict cognitive as well as noncognitive qualities (McDaniel & Nguyen, 2001). SJTs include different scenarios that are representative for problematic situations in a specific context. The scenarios are mostly presented either in a paper-pencil or video-based format. Applicants are provided with different answer alternatives from which one is considered being best and one being worst. They have to indicate what they are most and least likely to do. A differentiation can be made between knowledge based and behavioral based SJTs. Whereas knowledge based SJTs are stronger related with cognitive abilities, behavioral based SJTs are better predictors of personality (McDaniel, Hartman, Whetzel, & Grubb, 2007). Oswald and colleagues (2004) have shown that when predicting absenteeism and self-rated performance, SJTs provide incremental validity above and beyond traditional predictors and personality. Researchers comparing the predictive validity of video-based and written SJTs focusing on interpersonal performance have suggested that video-based SJTs are better predictors of interpersonally oriented criteria than their written counterparts (Lievens & Sackett, 2006). Furthermore, a video-based SJT measuring students' procedural knowledge about interpersonal skills has been shown to be predictive of internship performance (measured 7 years later) and job performance (measured 9 years later; Lievens & Sackett, 2012). Empirical findings have provided evidence that even in an actual application situation SJTs reduce gender related subgroup differences (Lievens & Coetsier, 2002). However, despite the aforementioned advantages researchers have also claimed that especially behavioral based SJTs are prone to faking (Nguyen, Biderman, & McDaniel, 2005). Thus, the likelihood that applicants "fake good" on a behavioral based SJT especially in a high stakes testing situation might limit the utility as a predictor of noncognitive academic performance.

Biographical data (biodata) measures are another valid option to predict future academic behavior (Oswald et al., 2004). They include structured questions about applicants' background, life history and previous experiences.

This method is based on the idea that past behavior is a valid predictor of future behavior (Wernimont & Campbell, 1986). Whether the required information is directly verifiable (e.g., “In how many different languages besides English can you converse well enough to order a meal?” Schmitt, et al., 2003) or relies on subjective judgments (e.g., “Think about the last several times you have had to learn new facts or concepts about something. How much did you tend to learn?” Oswald et al., 2004) depends on the design of the instrument. Empirical findings have shown that biodata measures explain incremental variance above and beyond traditional predictors when predicting performance in terms of grades and citizenship behavior (Schmitt et al., 2009). Furthermore, research conducted in the work context has shown that biodata measures explain incremental variance above and beyond GMA (Mount, Witt, & Barrick, 2000) and personality (McManus & Kelly, 1999) when predicting performance. Even in predicting citizenship behavior, biodata measures explain more variance than is explained by personality alone (McManus & Kelly, 1999). With respect to counterproductive behavior, biodata measures might be theoretically effective as past counterproductive behavior is a good predictor of future counterproductive behavior (Harding, Carpenter, Finelli, & Passow, 2004). However, especially with respect to counterproductive behavior it is difficult to verify the accuracy of applicants’ reactions. Thus, assuming that applicants are likely to respond in a socially desirable manner when asked about previous deviant behavior, biodata measures might be less appropriate to predict counterproductive behavior.

When it comes to informant ratings, letters of recommendations are widely used among universities (Liu, Minsky, Ling, & Kyllonen, 2009). Letters of recommendations are provided by former supervisors and include information about applicants’ past performance and noncognitive qualifications (McCarthy & Goffin, 2001). A recent meta-analysis has shown that letters of recommendation share a small but meaningful correlation with degree attainment, but not with grades or faculty ratings (Kuncel, Kochevar, & Ones, 2014). Asking academic and applied professionals about the importance they place on the content of letters of recommendations when taking selection decisions, academic professionals have reported giving them relatively more weight than practitioners (Nicklin & Roch, 2009). However, both groups appear to place little weight in absolute terms questioning the utility of letters of recommendations (Nicklin & Roch, 2009). This might be due to several shortcomings associated with the use letters of recommendations such as limited reliability, possible misinterpretations due to cultural or language differences and lack of comparability (Liu, et al., 2009; Range et al., 1991).

Taken together, previous research on methods measuring noncognitive factors has provided crucial evidence for the importance of noncognitive assessments. However, there are certain problems associated with the aforementioned predictors reducing their usefulness in practice. Therefore, in the following I will introduce the ETS Personal Potential Index® (PPI) an instrument that has been recently developed to overcome typical shortcomings such as social desirable responding and missing comparability. By combining aspects of commonly used traditional letters of recommendation with those of standardized rating forms, the PPI predicts students' noncognitive academic potential. It systematically assesses six noncognitive predictors of study success: Knowledge and creativity, communication skills, teamwork performance, resilience, planning and organization, ethics and integrity. The PPI has been developed to collect information from former supervisors about graduate students' noncognitive academic potential by using a web-based evaluation system. Just as letters of recommendation, the PPI is not based on self-ratings, but on supervisor ratings and is therefore not affected by social desirable responding and faking. In contrast to letters of recommendations, the PPI consists of a pre-defined set of items and is therefore highly standardized. Thus, taken together the PPI might be a promising instrument in order to obtain a more complete picture of the applicant. However, empirical research investigating the predictive validity of the PPI is missing so far. Assuming that the PPI might be especially useful to predict an extended criterion domain this dissertation aims to investigate the validity of the PPI when predicting citizenship and counterproductive behavior in the academic context.

Table 1.2 Predictive Validity Evidence for Noncognitive Predictors of Study Success

Method	Study	Empirical findings
Self-assessments	Poropat (2009)	CO, AG, and OP share a significant positive relationship with GPA
	Gore, et al., (2012)	CO shares a significant positive relationship with UCB
	Marcus, et al., (2007)	CO, EX, and HH share a significant negative relationship with CAB
	Robbins et al., (2004)	Academic self-efficacy and achievement motivation share a positive significant relationship with GPA above and beyond traditional predictors
Situational judgment tests (SJT)	Richardson, et al., (2012)	Grade goal shares a positive significant relationship with GPA above and beyond traditional predictors
	Credé & Kuncel (2008)	Study habits share a positive significant relationship with GPA above and beyond traditional predictors
	Lievens & Coetsier (2002)	SJTs provide incremental validity above and beyond traditional predictors when predicting GPA
	Lievens & Sackett (2012)	SJTs measuring procedural knowledge about interpersonal skills predict internship and job performance above and beyond cognitive predictors
Biodata measures	Oswald et al., (2004)	SJTs provide incremental validity above and beyond traditional predictors and personality when predicting absenteeism and self-rated performance
	Schmitt et al., (2009)	SJTs provide incremental validity above and beyond traditional predictors when predicting GPA and absenteeism
	Oswald et al., (2004)	Biodata measures provide incremental validity above and beyond traditional predictors when predicting GPA and absenteeism
	Schmitt et al., (2009)	Biodata measures provide incremental validity above and beyond traditional predictors when predicting GPA and UCB
Other ratings	Kuncel, et al., (2014)	Letters of recommendation share a meaningful correlation with degree attainment, but not with GPA or faculty ratings

Note. GPA = Graduate grade point average; UCB = University citizenship behavior; CAB = Counterproductive academic behavior; CO = Conscientiousness; AG = Agreeableness; OP = Openness to Experience; EX = Extraversion; HH = Honesty-humility

Other Antecedents of Counterproductive Academic Behavior

The consequences of counterproductive behavior are especially severe from universities' perspective. Thus, empirical research investigating the mechanisms influencing the occurrence of counterproductive behavior is required. A recently published study (Krishnakumar & Robinson, 2015) provides evidence for the negative relationship between trait mindfulness and counterproductive work behavior. Mindfulness can be described as awareness that emerges through paying attention on purpose, in the present moment, and nonjudgmentally to the unfolding of experience moment by moment (Kabat-Zinn, 2003). Originally known from Buddhist traditions as one of the seven steps towards enlightenment, the concept first obtained attention in clinical psychology due to the effectiveness of mindfulness-based interventions (e.g., mindfulness-based stress reduction [MBSR]: Kabat-Zinn, 1982; mindfulness-based cognitive therapy [MBCT]: Segal, Williams, & Teasdale, 2002). Interventions targeted to increase mindfulness have proven to be successful in decreasing symptoms and causes of health related issues (Baer, Fischer, & Huss, 2005; Bohlmeijer, Prenger, Taal, & Cuijpers, 2010; Hofmann, Sawyer, Witt, & Oh, 2010; Kabat-Zinn, 1982). The stream of literature, focusing on nonclinical populations, supports the positive effect of mindfulness on well-being indicators (Brown & Ryan, 2003; Caldwell, Harrison, Adams, Quin, & Greeson, 2010; Collard, Avny, & Boniwell, 2008; Hülshager, Alberts, Feinholdt, & Lang, 2013; Regehr, Glancy, & Pitts, 2013) and performance-related behaviors (Dane, 2011; Glomb, Duffy, Bono, & Yang, 2011; Shao & Skarlicki, 2009).

The relation between mindfulness and counterproductive behavior is mediated by negative feelings such as irritation and anger (Krishnakumar & Robinson, 2015). Thus, people scoring high on trait mindfulness were less prone to negative feelings which in turn reduced the occurrence of counterproductive behaviors. The idea that people with high mindfulness levels are less likely to have negative feelings is in line with findings on the neurobiological level. In an fMRI study, researchers compared two groups with respect to activation in the prefrontal cortex when confronted with unpleasant pictures (Lutz et al., 2014). The experimental group received a short mindfulness instruction, whereas the control group did not receive any emotion regulation strategies. When confronted with the negative pictures, participants in the mindfulness condition showed reduced activity in brain regions associated with emotion processing. Thus, it might be that the participants in the mindfulness condition perceived less arousal in response to the negative pictures. Taken together, students' individual mindfulness level might influence how they perceive certain situations. Whereas a more mindful student might feel positively challenged by an upcoming exam period (medium arousal) a less mindful student might experience the same situation as threatening and stressful (too much arousal).

How students finally react to their perception and whether it will be translated into counterproductive academic behavior, however, might depend on stable personality characteristics. Especially conscientiousness and honesty-humility are well-established predictors of counterproductive behavior (de Vries, et al., 2011; Marcus et al., 2007). A conscientious student might try to engage in constructive problem solving activities in order to handle the work load, whereas a less conscientious student exhibits counterproductive behavior to meet the study-related requirements. Thus, especially those students who are at risk of showing counterproductive academic behavior due to a low conscientiousness and honesty-humility level might profit from a high mindfulness level as they experience situations less stressful or threatening than their less mindful counterparts. Therefore, when investigating the link between trait mindfulness and counterproductive academic behavior, I also take the moderating role of conscientiousness and honesty-humility into account.

Dissertation Outline

Drawing on theory developed in the context of I/O psychology, the overarching goal of this dissertation is to provide new insights into the topic of graduate students' academic performance. The studies presented in this dissertation are based on data that has been collected in a comprehensive research project with a longitudinal set-up. A short overview of the project and the different steps of data collection is provided before the content of the different chapters is described in more detail. In doing so, I will focus on those parts of the project that are relevant for the present dissertation.

Between September and November 2011, international master students from five different faculties (School of Business and Economics, Faculty of Neuroscience, Faculty of Law, Faculty of Health, Medicine and Life Science, and Faculty of Arts and Social Science) were invited to participate in this study and do the newly launched GRE General test. Various communication channels were used to inform students about the possibility to participate in the project. Participation was voluntary. However, the benefits of participating in the project were emphasized: (1) the possibility to increase their international employability with a strengthened CV including GRE and PPI results, (2) the possibility to gain experience with an approved, internationally known assessment tool, which may help to improve their performance when they apply for jobs or postgraduate education programs in future, and (3) the opportunity to participate in a lottery including ten iPads

Before participants started working on the revised version of the GRE, they were informed about the different phases of the project and signed an informed consent document. After they finished the GRE they received detailed instructions how to create a PPI account on the ETS website in order to invite former supervisors to provide an evaluation. They were asked to fill in email addresses of two to five former supervisors. Finally, participants had to fill in a questionnaire including questions about demographic characteristics such as gender, age, country of origin, socio-economic background and U-GPA and a personality and trait mindfulness inventory.

The overall sample consisted of 282 participants with an average age of 23.55 ($SD = 2.36$). 155 participants were female and 127 were male. Overall, participants owned citizenship of 37 different countries. This is in accordance with the general student population enrolled at Maastricht University which is known for its diverse student body. A total of 35% of the participants were Dutch and 29% were German. The remaining 36% possessed citizenship of numerous other countries. For the Faculty of Psychology and Neuroscience (FPN) and the School of Business and Economics (SBE), U-GPAs were available for a larger cohort of master students who received their U-GPAs from Maastricht University allowing us to compare master students who participated ($n = 22$ for SBE; $n = 13$ for FPN) and master students who did not participate ($n = 128$, $n = 116$) with respect to their previous academic performance. Results of this comparison revealed that former undergraduate students who participated had U-GPAs that were similar to U-GPAs of former undergraduate students who did not participate at both FPN, $M_{\text{participants}} = 7.54$ ($SD = .33$), $M_{\text{non-participants}} = 7.52$ ($SD = .59$), $t = .11$, $p = .91$, and SBE, $M_{\text{participants}} = 7.23$ ($SD = .60$), $M_{\text{non-participants}} = 7.21$ ($SD = .57$), $t = .12$, $p = .91$. The final sample, for which PPI evaluations were available, consisted of 115 graduate students (female, $n=61$, male, $n=54$; age = $M: 23.85$, $SD: 3.00$)

Approximately three months after collection of predictor data, criterion data collection took place. To that end, participants were asked to complete a questionnaire assessing their contextual performance (university citizenship behavior and counterproductive academic behavior). Two-hundred sixty-one participants (response rate 93%) filled in the questionnaire.

Two years after the GRE data collection took place, faculties' student registry offices provided information on G-GPA, master thesis grades, time to complete, and degree attainment within scheduled time frame. For 236 participants (84%) the information was available. The remaining participants had not yet finished their master program or dropped out of the program. To provide an overview about the intercorrelations between all study variables Table 1.3 is provided.

Table 1.3 Correlations Between All Study Variables

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
	<i>n</i>	Correlations																
1. GRE-V	278	—																
2. GRE-Q	278	.41**	—															
3. GRE-AW	279	.42**	.30*	—														
4. PPI	115	.04	.01	.15	—													
5. MIND	281	-.03	-.08	-.01	.00	—												
6. U-GPA	257	.13*	.10	.20**	.32**	.07	—											
7. SES	279	.05	.05	.02	.10	.10	.16*	—										
8. OP	281	.19**	-.07	.17**	.09	.04	.16*	.23**	—									
9. CO	281	-.01	-.05	.06	.04	.35**	.20**	-.02	-.01	—								
10. AG	281	-.03	-.07	-.06	.08	.20**	-.01	-.04	.08	.01	—							
11. EX	281	-.07	-.01	.01	.08	.24**	-.06	.06	.17**	.11	.09	—						
12. EM	281	-.17**	-.16*	-.05	.02	-.11	.19**	-.06	-.03	.08	-.01	-.14*	—					
13. HO	281	.12*	-.14*	.09	.13	.14*	.13*	.04	.17**	.15*	.31**	-.11	.15*	—				
14. UCB	261	.03	.02	.11	.23*	.14*	.13*	.09	.25**	.26**	.18**	.25**	-.02	.13*	—			
15. CAB	261	-.01	.07	-.11	-.29**	-.28**	-.17**	-.01	-.04	-.47**	-.14*	.04	-.17**	-.29**	-.20**	—		
16. G-GPA	236	.21**	.17**	.31**	.10	-.10	.22**	.19**	.06	.15*	-.02	.00	.24**	.04	.10	-.09	—	
17. Degree attainment	282	.03	.09	-.01	.12	.13*	.14*	.03	-.06	.15*	.09	.04	.17**	.06	.16*	-.27**	.20**	—

Note. GRE-V = GRE Verbal Reasoning (130-170); GRE-Q = GRE Quantitative Reasoning (130-170); GRE-AW = GRE Analytical Writing (0-6); PPI = PPI Overall Score (1-5); MIND = Mindfulness; U-GPA = Undergraduate grade point average (0-10); SES = Mean score of father's and mother's education (1-7); OP = Openness to experience (1-5); CO = Conscientiousness (1-5); AG = Agreeableness (1-5); EX = Extraversion (1-5); EM = Emotionality (1-5); HO = Honesty-Humility (1-5); UCB = University citizenship behavior (1-5); CAB = Counterproductive academic behavior (1-5); G-GPA = Graduate grade point average (0-10); Degree attainment within scheduled time frame was coded 1 = finished in time, 0 = finished later than scheduled time frame.

* $p < .05$ ** $p < .01$.

Summing up, this dissertation aims to support universities facing the need to select well-suited graduate students with the potential to develop over the course of a master program in order to become valuable job market candidates by means of a longitudinal study. In doing so, Chapter 2 provides validity evidence for the GRE by focusing on two outcome measures that are indicative of study success: graduate grade point average (G-GPA) and degree attainment within scheduled time frame. I controlled the influence of students' U-GPA in order to make sure that the GRE explains additional variance above and beyond this widely used predictor of study success. Furthermore, socioeconomic status has been taken into account, as standardized tests are often criticized to be only predictive of students' social background. Results reveal that the Analytical Writing part of the GRE is a valid predictor of students' graduate grade point even if the influence U-GPA and socioeconomic status is held stable. Thus, it might be reasonable to use the GRE as predictor of students' G-GPA in the European context.

In Chapter 3 I gather systematic information on the predictive validity of the PPI by broadening the criterion domain and focusing on contextual aspects of study performance namely university citizenship and counterproductive academic behavior. Again variance explained by students' U-GPA is controlled for. Indeed, supervisor ratings of students' academic potential are predictive of self-rated citizenship and counterproductive behavior measured after a three months' time lag. Thus, in case universities are specifically interested in the prediction of students' behavior going beyond task performance operationalized in terms of grades, the PPI might be a promising tool. The findings of this study support the importance of noncognitive correlates in admission decision making and underscore the need to distinguish between different aspects of performance when investigating the psychometric properties of noncognitive correlates.

In Chapter 4 I establish the relationship between trait mindfulness and counterproductive academic behavior. Assuming that students' personality alters the relationship between mindfulness and counterproductive behavior, I also investigate the moderating role of conscientiousness and honesty-humility the two strongest predictors of counterproductive behavior in the academic context. Results reveal that students' with high mindfulness levels report to engage less in counterproductive behavior after a three months time lag. The relation is moderated by conscientiousness and honesty-humility in that sense that the high propensity for careless and dishonest students to engage in counterproductive behavior is reduced when they score high on dispositional mindfulness. Thus, in case universities aim to reduce the occurrence of counterproductive academic behavior by training students to become more mindful, this might be especially fruitful for students with a high dispositional tendency towards undependable and dishonest behavior.

Finally, in Chapter 5, I summarize the main findings of each study and provide practical and theoretical implications. Furthermore, the limitations of this dissertation and corresponding ideas for future research are discussed.

Chapter 2

Graduate Student Selection: GRE,
Socioeconomic Status and Undergraduate
Grade Point Average as Predictors of Study
Success in a Western European University

Published as:

Schwager, I. T. L., Hülshager U. R., Bridgeman, B., & Lang, J. W. B. (2015). Graduate Student Selection: GRE, Socioeconomic Status and Undergraduate Grade Point Average as Predictors of Study Success in a Western European University. *International Journal of Selection and Assessment*, 23, 71-79.

Abstract

Graduate students' mobility has increased within Europe. Yet, empirical evidence on the validity of standardized admission tests in Europe is still scarce. The aim of the present study was therefore to investigate the incremental validity of the GRE® revised General Test above undergraduate grade-point average (U-GPA) by focusing on a multinational sample of master students (N = 282) enrolled at a Dutch university. Results indicated that the Analytical Writing part of the GRE predicts graduate grade point average (G-GPA) above and beyond U-GPA. Furthermore, results suggest that the relationship is independent of students' socioeconomic status as indexed by parental education. Implications for graduate student selection in a European context are discussed.

Keywords: admission testing, academic performance, Graduate Record Examination, undergraduate grade point average, socioeconomic status

European universities increasingly try to attract international students especially on the graduate level. In doing so, they intend to improve the quality of higher education, to encourage knowledge exchange, and to stimulate the national labor market (Becker & Kolster, 2012). Furthermore, national policy makers seek to strengthen the so-called pull factors which determine universities' attractiveness for foreign students (Mazzarol & Soutar, 2002). More and more graduate students take the opportunity to study abroad in order to gain international experience and to boost their employability. As a consequence, student mobility has increased considerably over recent years in Europe. With increasing student mobility, universities face the necessity to use cost effective, fair, and valid methods of student selection in order to evaluate such a diverse body of applicants and to enhance the likelihood of admitting students that are well suited and truly have the potential to contribute to the aforementioned objectives.

Standardized tests such as the Graduate Record Examination (GRE) are a reasonable alternative to the frequent use of undergraduate grade-point average (U-GPA) in selecting graduate students (Kuncel, & Hezlett, 2007). In contrast to U-GPA, test scores resulting from the three GRE parts measuring Verbal, Quantitative, and Analytical abilities can be easily compared among large numbers of applicants with different educational and national backgrounds. Recent meta-analyses have provided evidence for the predictive validity of such standardized admission tests (Kuncel, Credé, & Thomas, 2007; Kuncel, Hezlett, & Ones, 2001; Kuncel, Wee, Serafin, & Hezlett, 2010; Powers, 2004). Authors of the largest meta-analysis including 1,521 studies (Kuncel, et al., 2001) provided evidence for the validity of the GRE General Test in predicting G-GPA, faculty ratings of performance, and comprehensive examination scores.

However, the meta-analyses cited above (Goldberg & Alliger, 1992; Kuncel, et al., 2001; Kuncel, et al., 2010; Morrison & Morrison, 1995) are based predominantly on criterion data from the United States. Criterion data for the European context is limited. We are only aware of research that follows up the criterion validity of tests used for undergraduate selection up to the graduate level (Hell, Trapmann, & Schuler, 2007; Lievens, Buyse, & Sackett, 2005). However, studies providing criterion data for standardized tests specifically aimed at selecting graduate students into European institutions after they complete their undergraduate degree are scarce. Yet, mobility of graduate students in Europe is greatly increasing in recent years (e.g. Statistics Netherlands, 2010) and consequently graduate student selection is a major policy problem in many European countries.

Validity evidence obtained in the US context may not be generalizeable to the European context without further testing, as cultural differences and differences in educational systems may affect validities of selection methods (Hülshager, Maier, & Stumpp, 2007; Salgado & Anderson, 2002). For instance, the predictive validity of the GRE may be affected by different grading traditions: European education systems differ from the US system with respect to grading scheme and grading severity. The Netherlands, for example, has a grading culture which is much stricter than the US grading culture as the highest possible grade (a 10 on a 1-10 scale) is hardly ever awarded in the Netherlands (Nuffic, 2006). Furthermore, educational systems are characterized by their degree of differentiation and tracking at the secondary school level. In stratified school systems students are sorted into different types of schools preparing them for different educational and occupational trajectories according to their level of achievement. In less stratified educational systems comprehensive secondary schools are common and ability-based tracking is less widespread (Kerckhoff, 2001).

Stratification has an influence on the homogeneity of the student body in different countries which in turn might influence the predictive validity of standardized tests.

The first goal of the present study is to investigate the validity of the GRE to predict study success at the graduate level. This article reports the first GRE validation study conducted in an European country that we are aware of. To measure study success, we not only consider graduate grade point average (G-GPA), but also degree attainment within scheduled time frame. Whereas G-GPA is one of the most widely used measures of graduate study success (see Kuncel et al., 2001), degree attainment within the scheduled time frame defined in the curriculum is of special interest because universities need to estimate the resources which need to be allocated on every single student based on predefined study duration. In case many students exceed this time span, additional resources are needed which in turn can have a negative impact on universities' capacities. In analyzing GRE validities for these outcomes, we control for U-GPA in order to assure that the GRE is a valid predictor of study success above and beyond U-GPA. U-GPA is widely used as a selection instrument in Europe. From a utility perspective it is therefore important to investigate the incremental validity of the GRE above this well-established predictor to make sure that an added value is obtained when using the GRE for student selection.

A second goal of the present paper is to investigate whether the validity of the GRE in predicting criteria of study success at a European university is not an artifact of socioeconomic status (SES). Critics of post-secondary admissions tests like the GRE have asserted that relationships between standardized admissions tests and criteria of study success are only an artifact of parental SES in terms of parents' earnings or education (e.g., Crosby, Iyer, Clayton, & Downing, 2003; Kohn, 2001; Zwick, 2004). This assertion rests on the idea that SES influences both grades and test results in admissions tests, such that any association found between admissions tests and grades could be explained by SES and therefore not be indicative of the predictive validity of the test itself. Notably this criticism also includes the idea that a positive association between SES and test scores does not reflect a true higher standing of high SES students on what admissions tests measure but simply higher familiarity with test-taking. Similarly, the criticism implies that any association found between SES and study grades is not caused by higher true academic performance for students high in SES but asserts that it is an artifact of grading being biased in favor of high SES students (for a detailed discussion see Sackett, Bornemann, & Connelly 2008; Sackett, Kuncel, Arneson, Cooper, & Waters, 2009). Following this criticism, any relation found between test scores and criteria of study success would be an artifact of SES and should therefore disappear when one controls for parental SES. Sackett and colleagues (2009) tested this idea in a comprehensive study including various large-scale US datasets on the SAT (formerly Scholastic Aptitude Test). They found that although SES was related to SAT test scores, the SAT-grade relationship was only marginally reduced when the influence of SES was controlled for. The present study sets out to replicate Sackett et al.'s (2009) findings in a European sample and using a graduate admission test (GRE) rather than an undergraduate admission test (SAT).

Method

Participants and Procedure

Two-hundred eighty-two international students (female = 155; male = 127; age: $M = 23.55$, $SD = 2.36$) enrolled at five faculties (Business [$n = 83$], Psychology [$n = 125$], Law [$n = 49$], Medicine [$n = 16$], and Arts and Social Science [$n = 9$]) at a Dutch University volunteered for the project at the beginning of their master program. They indicated that they originated from 22 European (e.g., Netherlands, Germany, Bulgaria, Poland, Lithuania, Belgium, Spain), and 15 non-European countries (e.g., China, Turkey, Ukraine, Canada, Mexico, Colombia, Vietnam). Furthermore, they received their bachelor degrees from 35 different countries and were enrolled in English-speaking master programs. This is in line with the assumption that Dutch universities have a strong international focus due to the central location within Europe. Furthermore, numerous master's degree programs are taught in English in order to reduce potential language barriers (Becker & Kolster, 2012) and to strengthen the attractiveness of Dutch universities for foreign students (Mazzarol & Soutar, 2002). The majority of participating students were enrolled in 1-year master programs; only six participants were enrolled in a 2-year research master program.

The study had a longitudinal set-up, with GRE testing taking place at the beginning of participants' master studies and criterion data collection taking place approximately two years later. At time 1, master students participated in a four hour testing session, in which they completed the revised (now current) version of the GRE, and a questionnaire including demographic information (age, gender, faculty affiliation, SES, U-GPA).

Criterion data was collected at time 2, two years after completion of the GRE. Specifically, information on students' graduate grade point average (G-GPA) and degree attainment was obtained from student registry offices. At time 2, G-GPAs were available for 236 participants (84%). The remaining participants had not finished their master program until this date. For our second criterion, degree attainment within scheduled time frame, the entire sample ($N = 282$) has been taken into account.

Measures

Graduate Record Examination (GRE) revised General Test. We used the revised GRE General Test, consisting of three parts: GRE-V (Verbal Reasoning), GRE-Q (Quantitative Reasoning), and GRE-AW (Analytical Writing). Whereas GRE-V measures the ability to analyze and evaluate written material and GRE-Q measures basic mathematical skills and the understanding of mathematical concepts, GRE-AW measures critical thinking and analytical writing skills. GRE-V and GRE-Q scores are reported on a scale from 130 and 170 in 1-point increments. For GRE-AW the score range is from 0-6 in half-point increments.

Socioeconomic status (SES). Participants were asked to indicate the highest educational degree of their parents respectively. They had to choose one of six options: 1 = No school degree; 2 = Basic school degree; 3 = Advanced school degree; 4 = Vocational degree; 5 = Basic university degree; 6 = Advanced university degree; 7 = Doctoral degree. A mean score for mothers' and fathers' education was calculated.

Undergraduate grade point average (U-GPA). We asked participants to indicate their undergraduate grade point average. Since the majority of the sample provided Dutch grades, we converted the remaining grades into the Dutch grading format (0 = lowest grade to 10 = highest grade). Given that 10s are hardly ever provided in the Dutch grading system (Nuffic, 2006; 6= 33-37%, 7= 33-37%, 8= 22-26%, 9= 6%, 10= 1%), the highest attainable grades in other national systems were converted into a 9.0.

Graduate grade point average (G-GPA) and degree attainment within scheduled time frame. Faculties' student registry offices provided us with participant's G-GPA and start and graduation date. In case students attained their degree within the scheduled time frame (in most cases 1 year), the outcome variable was coded with 1. In case it took students longer to finish their master program the variable was coded with 0.

Data Analysis

To investigate GRE's predictive validity, we calculated zero-order correlations between GRE-V, GRE-Q, and GRE-AW and the two outcome measures (G-GPA and degree attainment within scheduled time frame). Grades provided over the course of several months by different supervisors affiliated with different faculties are prone to rating errors. One may therefore wish to correct for the fact that measurement error in the criterion may lead to lower observed correlations between GRE and G-GPA, thereby leading to an underestimation of the true validity of the GRE. In accordance with previous studies investigating the predictive validity of standardized admission tests (Kuncel, et al., 2007; Kuncel, et al., 2001; Kuncel, et al., 2010) we therefore corrected the observed GRE-G-GPA correlations for attenuation due to measurement error in the criterion (Hunter & Schmidt, 2004). In order to get an estimation of the reliability of G-GPA, we computed the reliability coefficient between the mean grade students received for their course work and the grade they received for their master thesis project. The course work grade is the average grade students obtain for four different courses providing general knowledge in their corresponding master specialization. In order to finish their studies, students have to conduct an empirical research project based on which they write their master thesis. The course work and the master thesis project are equally important with respect to earned credits and sum up into the final average grade. The raw correlation of .38 ($p < .01$) between the practical part and the research based part resulted in a reliability of .54 for the G-GPA estimate. Accordingly, we divided the observed correlations between the independent variables and G-GPA by the square root of the reliability coefficient (Hunter & Schmidt, 2004) in order to obtain an estimate of the operational validity of the GRE. For degree attainment within scheduled time frame no correction was made as this is an objective criterion that is not affected by measurement error.

In order to examine the influence of parental SES and the incremental validity of the GRE above and beyond U-GPA when predicting G-GPA, we used hierarchical regression analysis. For G-GPA as dependent variable, we ran two regression analyses – one based on the raw data and one based on the correlation matrix including the correlations corrected for criterion unreliability. Since degree attainment within scheduled time frame was a dichotomous variable, we used a logistic regression analysis to analyze the data. In Step 1, we controlled for SES, in Step 2 we entered U-GPA, and in Step 3 we included GRE-V, GRE-Q, and GRE-AW in both regression analyses.

Results

Table 2.1 provides an overview about the zero-order correlations between the predictor-criterion combinations in order to provide information about the predictive validity of the GRE. Considering the entire sample, GRE-V, GRE-Q and GRE-AW were all significantly related to subsequent G-GPA. In contrast, none of the three GRE subscales were significantly related to whether students attained their master degree within the scheduled time frame.

Table 2.1 Intercorrelations Between GRE Results, U-GPA, SES and Outcome Variables

		Correlations								
	<i>n</i>	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7
Complete sample										
1. GRE-V	278	146.80	6.91	—						
2. GRE-Q	278	147.71	7.56	.40**	—					
3. GRE-AW	279	3.11	0.63	.42**	.30**	—				
4. SES	279	4.76	1.18	.05	.05	.02	—			
5. U-GPA	257	7.61	0.67	.13*	.10	.20**	.16*	—		
6. G-GPA	236	7.53	0.53	.21** (.28)	.17** (.23)	.31** (.42)	.19** (.26)	.22** (.30)	—	
7. Degree attainment	282	0.56	0.50	.03	.09	-.01	.03	.14*	.20**	—
Psychology student sample										
1. GRE-V	124	145.90	6.47	—						
2. GRE-Q	124	144.83	6.02	.43**	—					
3. GRE-AW	125	3.03	0.71	.46**	.30**	—				
4. SES	125	4.81	1.22	.03	.05	.09	—			
5. U-GPA	115	7.55	0.63	.27**	.12	.37**	.08	—		
6. G-GPA	97	7.66	0.50	.25* (.40)	.24* (.38)	.42** (.67)	.14 (.22)	.45** (.72)	—	
7. Degree attainment	126	0.62	0.49	.05	.14	.02	-.06	.21*	.02	—
Business student sample										
1. GRE-V	83	147.83	6.88	—						
2. GRE-Q	83	153.19	7.46	.42**	—					
3. GRE-AW	83	3.18	0.52	.34**	.23*	—				
4. SES	83	4.75	1.14	.13	.26*	-.14	—			
5. U-GPA	77	7.59	0.66	.24*	.51**	.17	.26*	—		
6. G-GPA	79	7.48	0.54	.29* (.37)	.24* (.30)	.28* (.35)	.25* (.32)	.30* (.38)	—	
7. Degree attainment	83	0.52	0.50	.13	.26*	.08	.07	.27*	.26*	—

Note. GRE-V = GRE Verbal Reasoning; GRE-Q = GRE Quantitative Reasoning; GRE-AW = GRE Analytical Writing; SES = Mean score of father's and mother's education; U-GPA = Undergraduate grade point average; G-GPA = Graduate Grade Point Average. Degree attainment within scheduled time frame was coded 1 = finished in time; 0 = finished later than scheduled time frame. Correlations corrected for criterion unreliability are presented in parentheses.

* $p < .05$ ** $p < .01$

Table 2.2 provides the results for the incremental validity of the GRE above and beyond U-GPA when G-GPA functioned as dependent variables. Considering results based on uncorrected correlations, findings reveal that although U-GPA was a valid predictor of G-GPA ($\Delta R^2 = .03, p < .05$), the three GRE parts added a statistically significant increment to the prediction of G-GPA ($\Delta R^2 = .09, p < .001$). Especially GRE-AW was a statistically significant predictor of G-GPA ($\beta = .24, p < .05$). Table 2.3 depicts results for the logistic regression analysis with degree attainment within scheduled time frame as dependent variable. Compared to the three GRE parts, U-GPA emerged as a valid predictor of degree attainment ($B = .45, p < .05$). Conducting the same set of analyses based on correlations that were corrected for measurement error in the criterion point to even stronger GRE validities. When correcting for measurement error, the GRE parts explained an additional 15% of the variance in G-GPA above and beyond U-GPA.

In order to test whether the relationship between GRE and G-GPA was only an artifact of SES, we first inspected the zero-order correlations between GRE scales and parental SES (Table 2.1). The correlations were not statistically significant. This pattern of results indicated that GRE and SES shared little common variance. Results of the hierarchical regression analysis (Table 2.2) also revealed that GRE-AW still significantly predicted G-GPA ($\beta = .24, p < .01$) when the influence of parental SES was controlled. Accordingly, controlling for SES did not eliminate the predictive power of GRE-AW.

Table 2.2 Hierarchical Multiple Regression Analyses Predicting G-GPA from SES, UGPA, and GRE Results

Predictor	G-GPA	
	ΔR^2	β
Step 1	.02* (.07)	
SES		.15* (.26)
Step 2	.03* (.05)	
U-GPA		.18** (.20)
Step 3	.09*** (.15)	
GRE-V		.09 (.08)
GRE-Q		.03 (.07)
GRE-AW		.24** (.36)
Total R ²	.14*** (.27)	

Note. $n = 236$. G-GPA = Graduate grade point average; SES = Mean score of father's and mother's education; U-GPA = Undergraduate grade point average; GRE-V = GRE Verbal Reasoning; GRE-Q = GRE Quantitative Reasoning; GRE-AW = GRE Analytical Writing. Coefficients corrected for criterion unreliability are presented in parentheses.

* $p < .05$ ** $p < .01$ *** $p < .001$

Table 2.3 Logistic Regression Analysis Predicting Degree Attainment within Scheduled Time Frame from SES, U-GPA, and GRE

Predictor	Model 1	Model 2	Model 3
SES	.01	-.03	-.04
U-GPA		.43*	.45*
GRE-V			.01
GRE-Q			.02
GRE-AW			-.23
Model fit			
χ^2	0.02	4.82	6.75
<i>df</i>	1	2	5
Nagelkerke R^2	.00	.03	.04

Note. $N = 282$. SES = Mean score of father's and mother's education; U-GPA = Undergraduate grade point average; GRE-V = GRE Verbal Reasoning; GRE-Q = GRE Quantitative Reasoning; GRE-AW = GRE Analytical Writing. Degree attainment within scheduled time frame was coded 1 = finished in time; 0 = finished later than scheduled time frame.

* $p < .05$

Supplementary Analyses

One may wonder whether GRE validities were similar or differed across study areas. In order to provide some insights on this, we computed zero-order correlations for the two largest subsamples of Psychology and Business students. As can be seen from Table 2.1, correlations within the Psychology and Business students sample were similar to each other and similar to correlations found in the overall sample. One notable exception was that GRE-Q was significantly related to degree attainment within scheduled time frame in the Business student sample. In addition, we analyzed whether study area moderates the relationship between the three GRE parts and G-GPA (based on uncorrected correlations) and degree attainment within scheduled time frame. To do so, we formed two dummy variables differentiating between psychology, business studies, and the rest (law, medicine, and art and social sciences). As can be seen from Table 2.4 and 2.5, none of the interaction terms were significant indicating that GRE validities equal each other across faculties.

Table 2.4 Hierarchical Multiple Regression Analyses Predicting G-GPA

Predictor	G-GPA	
	ΔR^2	β
Step 1	.02*	
SES		.15*
Step 2	.03*	
U-GPA		.18*
Step 3	.09***	
GRE-V		.09
GRE-Q		.03
GRE-AW		.24**
Step 4	.08***	
Dummy 1: FPN and other (0) vs. SBE (1)		-.24**
Dummy 2: FPN and SBE (0) vs. Other (1)		-.29***
Step 5	.02	
GRE-VxDummy 1		.09
GRE-VxDummy 2		.03
GRE-QxDummy 1		-.13
GRE-QxDummy 2		.11
GRE-AWxDummy 1		.02
GRE-AWxDummy 2		-.01
Total R^2	.24	

Note. $n = 236$. G-GPA = Graduate grade point average; SES = Mean score of father's and mother's education; U-GPA = Undergraduate grade point average; GRE-V = GRE Verbal Reasoning; GRE-Q = GRE Quantitative Reasoning; GRE-AW = GRE Analytical Writing; FPN = Faculty of Psychology and Neuroscience; SBE = School of Business and Economics; Other = Faculty of Law, Faculty of Health, Medicine and Life Science, and Faculty of Arts and Social Science.

* $p < .05$ ** $p < .01$ *** $p < .001$

Table 2.5 Logistic Regression Analysis Predicting Degree Attainment

Predictor	Model 1	Model 2	Model 3	Model 4	Model 5
SES	.01	-.03	-.04	-.05	-.05
U-GPA		.43*	.45*	.47*	.45*
GRE-V			.01	.01	.00
GRE-Q			.02	.04	.05
GRE-AW			-.23	-.21	-.22
Dummy 1: FPN and other (0) vs. SBE (1)				-.85*	-.89*
Dummy 2: FPN and SBE (0) vs. Other (1)				-.67*	-.66
GRE-VxDummy 1					.03
GRE-VxDummy 2					.02
GRE-QxDummy 1					-.01
GRE-QxDummy 2					-.10
GRE-AxDummy 1					.12
GRE-AxDummy 2					-.09
Model fit					
χ^2	0.02	4.82	6.75	13.97	14.38
df	1	2	5	7	13
Nagelkerke R^2	.00	.03	.04	.07	.08

Note. $N = 282$. SES = Mean score of father's and mother's education; U-GPA = Undergraduate grade point average; GRE-V = GRE Verbal Reasoning; GRE-Q = GRE Quantitative Reasoning; GRE-AW = GRE Analytical Writing; FPN = Faculty of Psychology and Neuroscience; SBE = School of Business and Economics; Other = Faculty of Law, Faculty of Health, Medicine and Life Science, and Faculty of Arts and Social Science. Degree attainment within scheduled time frame was coded 1 = finished in time; 0 = finished later than scheduled time frame.

* $p < .05$

Discussion

The present study set out (a) to investigate the validity of the GRE for study success in a European university considering two different study criteria and (b) to test whether associations found between GRE test scores and study success criteria are independent of parental SES.

Regarding the first goal of the study, our findings provide evidence that the GRE is a valid predictor of graduate study success at a Dutch university. GRE-V, GRE-Q and GRE-AW measured at the beginning of the master program predicted students' G-GPA at the end of the master program. The finding that none of the other GRE scales predicted degree attainment within scheduled time frame might be explained by the fact that other, non-ability related factors may play a larger role in determining whether or not students complete their studies within the designated time frame, for instance financial resources, conscientiousness, motivation, or adaptability.

Overall, results found in the present study were consistent with previous meta-analyses investigating the predictive validity of the GRE in the US context (Kuncel et al., 2001; Kuncel et al., 2010). Our study results are similar to Kuncel and colleagues' (2001) meta-analytically derived average correlations with regard to GRE-V ($r = .21$, present study vs. $r_{obs} = .23$, Kuncel et al., 2001) and GRE-Q ($r = .17$ vs. $r_{obs} = .21$); GRE-AW was not included in the meta-analyses as it did not yet exist when the studies contributing to the meta-analyses were conducted. Instead of degree attainment which was defined as the successful completion of a degree program (Kuncel et al., 2001), we focused on degree attainment within the scheduled time frame in the present study. Although degree attainment and degree attainment within scheduled time frame are somewhat different, findings are similar with respect to GRE-Q ($r = .09$ vs. $r_{obs} = .14$). Furthermore, incremental validity evidence obtained in the present study demonstrates that the GRE predicts beyond what could be predicted from U-GPA alone. Especially GRE-AW provides supplemental information about future G-GPA.

In response to previous criticism that associations found between admissions test scores and criteria of study success are an artifact of SES (Crosby et al., 2003; Kohn, 2001; Zwick, 2004), a second goal of the present paper was to test whether GRE validity was independent of parental SES. Our findings revealed that (a) GRE test scores were not significantly related to parental SES and (b) that GRE-AW was still significantly related to G-GPA when SES was controlled in a hierarchical regression analysis. In order to provide further insights into the role of SES in GRE validities and to compare our findings to previous findings on the role of SES in validities of admission test at the undergraduate level (Sackett et al., 2009), we computed partial correlations (based on uncorrected correlations). Specifically, we partialled out the influence of SES from the correlations between GRE parts and G-GPA and compared these to the uncorrected zero-order correlations reported in Table 2.1.

Controlling for SES, GRE-G-GPA correlations remained virtually the same ($r_{\text{GRE-V-G-GPA}} = .21, p < .01$; $r_{\text{GRE-Q-G-GPA}} = .17, p < .01$; $r_{\text{GRE-AW-G-GPA}} = .30, p < .001$). The latter finding is consistent with Sackett and colleagues' (2009) finding from the United States that the SAT-grade relationship was only marginally reduced (the correlation dropped from .47 to .44) when the influence of SES was controlled. The present findings thus replicate Sackett and colleagues' finding at the graduate level. Yet, our study also revealed that GRE-scores were unrelated to parental SES and this finding differs from Sackett et al.'s finding that SAT-scores were substantially related to SES. Another interesting finding is that U-GPA and G-GPA were related to SES. This indicates that higher SES students tend to get higher grades at both the undergraduate level and the graduate level and that the association between U-GPA and G-GPA may – in part – be caused by parental SES. Providing that this finding can be replicated in future studies using different samples, it suggests that using U-GPA as a selection criterion may be less fair than using a standardized admission test like the GRE that was unrelated to SES in the present study.

Taken together, our results suggest that the GRE is a valid instrument that may meet the urgent need for objective selection instruments in a European context. Besides the predictive and incremental validity evidence obtained in the present study, the GRE meets other important requirements distinctive for a sound selection instrument. First, the GRE is objective as test results are easily interpreted and do not depend on individual evaluations which might be prone to biases. Consequently, the use of the GRE for student selection purposes is robust against legal claims that are often grounded on missing comparability. Furthermore, it is efficient from universities' perspective, but also from the perspective of master students. In case European universities would comprehensively agree upon using the GRE for master student selection purposes, students could use their GRE scores for several applications within the same, but also between different countries. Lastly, the GRE could even increase fairness of master students' selection. The present study demonstrated that parental SES has an influence on the relationship between U-GPA and subsequent study success, whereas the relationship between GRE scores and indicators of study success seems to be independent of factors that might be influenced by social class.

Practical Implications

Since 1999, countries within the European Union have worked toward increasing the comparability and compatibility of their higher education programs (Education, Audio-visual and Culture Executive Agency, 2010). One major goal of the Bologna process is to introduce a two-phase system with a bachelors and a master phase in all European countries (of which many have long used a one-phase system). A second goal of the Bologna process is to allow students to change between universities in different regions within European countries and also between European countries as they traverse from undergraduate to graduate education. As students increasingly use these new opportunities, European universities now increasingly need to select students with an undergraduate degree into their graduate programs. Currently, European universities commonly use applicants' undergraduate grade point average (U-GPA) to select graduate students. U-GPA is easily accessible, and its use is convenient for universities and applicants.

Especially due to the shortcomings of U-GPA the findings of the present study have important practical implications. In Germany, for example, a debate was going on whether bachelor students graduating from top universities are discriminated against students from less renowned universities when applying for a master program (Dörnfelder, 2010). This criticism is based on the assumption that bachelor students graduating from prestigious universities often get lower grades due to the high requirements, compared to students who obtained their bachelor degree from universities with lower standards. Furthermore, changes going along with the Bologna declaration have facilitated students' possibilities to follow a master program which is related, but not necessarily equal to the study field where the bachelor degree was obtained. As a result, bachelor students from different study fields (e.g., Business and Psychology) compete against each other for the same master programs (e.g., Marketing or Work and Organizational Psychology) although the two programs rely on different grading standards and are therefore difficult to compare. Accordingly, in order to guarantee a fair selection procedure the GRE is a promising supplement to U-GPA.

Limitations and Future Research

The present study is not without limitations. Participants were master students who have already started their master programs. They participated on a voluntary basis and did not practice beforehand. In this sense, the testing situation can be considered a low stakes situation which might have resulted in somewhat lower GRE scores compared to real applicants. This does not negate the validity evidence obtained in the present study. However, in order to define GRE cut-off scores for a European context results should be treated with caution.

Another limitation of this study is that we solely collected data at a Dutch university. Although the university is located close to the German and Belgium boarder and characterized by an international student body, future validity studies should be conducted in other European countries as well. Future studies may also benefit from including students from the science, technology, engineering and math (STEM) fields in order to get further information about the predictive validity of the GRE in these fields. It might be especially interesting to investigate whether the inclusion of students from STEM fields will change the predictive validity of GRE-AW as it might be that students more practiced in analytical writing (humanities and social sciences) show markedly different results from those in the STEM disciplines.

A third limitation exists with respect to the sample size of the present study. Researchers have called for large-scale studies and meta-analyses to investigate the characteristics and usefulness of standardized tests (Kuncel & Hezlett, 2010; Sackett, et al., 2008). Given that the present study is the first validity study conducted in a European context, we consider it as an initial step towards more comprehensive studies in this field. Large-scale studies would also allow analyzing subgroup differences. Although the adverse impact issue might be more about differences between students originating from different European countries than about ethnic minorities within countries it should not be neglected.

Finally, a limitation exists with respect to the measurement of SES. In order to measure students' SES, self-ratings of parental education were used which might be prone to biases. However, a study conducted by Loocker (1989) showed that the level of agreement between what students indicate as the highest educational degree of their parents and parents' self-evaluation was high (mother's education: $r = .85$; father's education: $r = .86$). This is in line with Sackett et al. (2009) who noted that students are less likely to provide erroneous information about their parents' education compared to alternative SES measures such as income.

Conclusion

The results provide evidence for the usefulness of the GRE in a European context. The GRE allows valid predictions of master students' study success in terms of graduate grade point average. It facilitates the decision making process of selection committees who have to decide upon students with various cultural and educational backgrounds.

Chapter 3

Supervisor Ratings of Students' Academic Potential as Predictors of Citizenship and Counterproductive Behavior

Published as:

Schwager, I. T. L., Hülshager U. R., Lang, J. W. B., Klieger, D. M., Bridgeman, B., & Wendler, C. (2014). Supervisor Ratings of Students' Academic Potential as Predictors of Citizenship and Counterproductive Behavior. *Learning and Individual Differences*, 35, 62-69.

Abstract

The educational psychology literature has recently discussed the implications of university citizenship behavior – activities that go beyond the primary student role – and counterproductive academic behavior – behaviors that run counter to academic values and objectives. This study investigated whether supervisor ratings of students' academic potential can serve as a valuable predictor of those criteria. The sample included 115 international students enrolled in a master's degree program at a university in the Netherlands. Results revealed that supervisor ratings of students' academic potential provided incremental validity above and beyond undergraduate grade point average (U-GPA) in predicting self-ratings of university citizenship and counterproductive academic behavior. The usefulness of supervisor ratings as a supplement to traditional predictors of study success, and implications for student selection are discussed.

Keywords: Academic Performance; University Citizenship Behavior; Counterproductive Academic Behavior; Noncognitive Predictors; Diversity

Harvard University recently suspended more than 60 students due to a cheating scandal (Pérez-Peña, 2013). The suspended students collaborated on a take home exam, in spite of strict instructions to work on it autonomously. Their behavior clearly had a negative impact on the reputation of the elite university, which ironically uses the Latin word for truth as its official slogan. Apparently, not all students who were admitted were strongly committed to Harvard's integrity principle. At the graduate level, cheating is a widespread issue as well (McCabe, Butterfield, & Trevino, 2006). This is especially alarming as graduate students are about to enter the job market where the consequences of unethical behavior are yet more severe compared to those in the academic context.

In efforts to take additional aspects of students' performance into account, researchers have recently suggested conceptualizing academic performance in broader terms (Lievens, Buyse, & Sackett, 2005; Meriac, 2012; Schmitt, Oswald, Friede, Imus, & Merritt, 2008; Sinha, Oswald, Imus, & Schmitt, 2011; Zettler, 2011). It has initially been acknowledged in the field of personnel selection, that performance is multidimensional, subsuming not only task but also contextual performance aspects such as citizenship and counterproductive behavior (Rotundo & Sackett, 2002; Viswesvaran & Ones, 2000). Task performance describes the fulfillment of tasks and responsibilities that are formally associated with a given job or role, whereas citizenship behavior refers to activities that are not necessarily related to the primary tasks of a given role, but that facilitate organizational effectiveness by contributing to the organizational, social and psychological environment (Borman & Motowild, 1997). While these two dimensions focus on activities that promote performance and benefit an organization, counterproductive behavior describes activities that stand counter to an organization's values and objectives and that harm the well-being of an organization (Rotundo & Sackett, 2002).

In a similar way, academic performance is multidimensional and entails not only task-related, but also university citizenship and counterproductive academic behavior (Meriac, 2012; Schmitt et al., 2008; Zettler, 2011). Task performance is certainly important in order to meet the formal requirements of a study program. University citizenship behavior, however, might also have a positive influence on students' academic success by stimulating cooperation between peer students, or faculty members, which in turn increases the likelihood of getting helpful guidance or feedback (Verduin & Clark, 1991). Counterproductive academic behavior, in contrast, interferes more directly with academic success, in cases where students do not attend a class or cheat on exams instead of learning the study materials (Credé & Niehorster, 2009). A major goal of universities is to prepare students for the job market (Morphew & Hartley, 2006), and to foster their employability (Poropat, 2011). Thus, the multidimensional conceptualization of performance plays an important role in the academic context as well. This is best represented by universities' mission statements that often highlight the importance of interpersonal and psychological aspects of study success (Oswald, Schmitt, Kim, Ramsay, & Gillespie, 2004). Furthermore, a recent meta-analysis provided evidence for the positive influence of education on future task performance, but also future citizenship and counterproductive work behavior (Ng & Feldman, 2009).

Although task performance is best represented by students' grades on exams or assignments, examples of university citizenship behavior include helping fellow students understand the subject matter, accepting inconvenient study conditions or engaging in voluntary activities. Counterproductive academic behavior includes behaviors such as copying from published sources, cheating on exams, stealing university property or bullying individuals affiliated with the university. Students sharing their knowledge to help other students provide better support to universities' educational missions than students who try to improve their own performance by cheating or stealing study materials. Accordingly, universities might benefit from selecting students based on their tendencies to engage in university citizenship and productive academic behavior, to increase the possibility of approving students who support universities' well-functioning instead of negatively affecting their reputation. In sum, higher levels of university citizenship behavior and reduced levels of counterproductive academic behavior have a positive influence on both, the students, but also the universities that benefit from a counterbalanced and committed student body.

In order to find an appropriate way to predict university citizenship and counterproductive academic behavior, we broaden the predictor domain and investigate the predictive validity of supervisor ratings of students' academic potential. Based on the two assumptions that past behavior predicts future behavior (Wernimont & Campbell, 1968) and that supervisory ratings are a reliable predictor of typical performance (Marcus, Goffin, Johnston, & Rohstein, 2007), we hypothesize that ratings of students' noncognitive academic potential provided by former supervisors are a valid predictor of future university citizenship and counterproductive academic behavior. Specifically, we investigate the predictive validity of the ETS® Personal Potential Index (ETS® PPI), a web-based evaluation system that uses former supervisors' ratings to systematically assess noncognitive aspects of academic potential. Given that the utility of a new selection approach is provided only in case additional variance is explained over existing and widely used selection instruments, we also investigate the incremental validity of the ETS PPI above and beyond undergraduate grade point average (U-GPA).

Noncognitive Predictors of Academic Performance and Study Success

Several researchers have already defined study success in broader terms and identified factors and instruments that might relate to this expanded conceptualization. For instance, in an early study, Enright and Gitomer (1989) interviewed university faculty members and identified seven competencies that are important for graduate study success: communication, creativity, explanation, motivation, planning, professionalism, and synthesis. In order to measure the seven dimensions, Enright and Gitomer (1989) proposed to use simulation tasks but they also critically discussed that using such simulation tasks are likely to be difficult to implement in large-scale student selection.

More recently, Oswald and colleagues (2004) systematically analyzed universities' mission statements and educational objectives in order to identify dimensions of undergraduate academic performance in broad sense. They defined twelve major dimensions which can be sub-categorized as intellectual behaviors (knowledge, learning, artistic), interpersonal behaviors (multicultural, leadership, interpersonal, citizenship), and intrapersonal behaviors (health, career, adaptability, perseverance, ethics). Based on their findings they developed a situational judgment test (SJT) and a biographical data (biodata) questionnaire. SJTs can be used to assess noncognitive skills, including descriptions of problematic situations individuals are instructed to solve. Answers are typically provided in a multiple response format or on a Likert scale. Biodata measures include structured questions about applicants' background and life history. SJTs and biodata measures have good predictive validities and even reduce subgroup differences in combination with cognitive predictors (McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001; Lievens & Coetsier, 2002; Oswald et al., 2004; Schmitt et al., 2009). However, the influence of coaching and faking is still an issue that needs to be taken into account when using these methods in practice (Ramsay, Schmitt, Oswald, Kim, & Gillespie, 2006).

Letters of recommendations are another option to assess a broader conceptualization of performance while reducing the effects of social desirable responding. Among universities they are widely used (Liu, Minsky, Ling, & Kyllonen, 2009). They are provided by former supervisors and include useful insights about students' knowledge, skills and personality (McCarthy & Goffin, 2001). Because they are based on other-ratings, they are less vulnerable to socially-desirable responding and cheating. A notable shortcoming of letters of recommendations is, however, that they suffer from typically weak psychometric properties (Range et al., 1991). As no guidelines exist for writing letters of recommendation, they are all written differently making it difficult to compare between applicants, and also to get a coherent picture of one applicant providing several letters. Standardization has been described as a successful means to improve the reliability and validity of letters of recommendation (Kim & Kyllonen, 2008; Kyllonen & Kim, 2005; McCarthy & Goffin, 2001; Nicklin & Roch, 2009; Schmitt, 2012) and faculty members and administrators prefer standardized letters of recommendation over traditional letters (Walters, Kyllonen, & Plante, 2006).

The ETS PPI is an instrument that combines advantages of traditional letters of recommendation (insights from former supervisors) with advantages of standardized rating forms (better psychometric properties). In contrast to traditional letters of recommendation which have no predetermined format or content, the ETS PPI measures students' academic potential with a set of items targeting six dimensions, thereby assessing the same content across applicants in a standardized way. Importantly, the ETS PPI relies on ratings from multiple (up to 5) supervisors, thereby enhancing reliability and validity. Previous research has documented that ratings from single other-raters are, to some extent, idiosyncratic which reduces reliability and attenuates correlations with outcome variables (Connelly & Ones, 2010). Using a composite score that is based on multiple raters is a way to overcome rater idiosyncrasies and increase reliability and validity (Connelly & Ones, 2010).

The six ETS PPI dimensions (see Table 3.1) are the result of interviews with deans and faculty members working at numerous universities (Briel et al., 2000), and a more in-depth follow up interview phase (Walpole, Burton, Kanyi, & Jackenthal, 2001). Walpole and colleagues (2001) asked faculty members to identify variables they felt were predictive of graduate school success and that could be used as part of the admissions process. A number of noncognitive variables were ranked highly (e.g., persistence, tenacity, collegiality, communication, and enthusiasm) and functioned as basis for the ETS PPI. Except for the teamwork dimension, all ETS PPI dimensions can either be found in the taxonomy of Enright and Gitomer (1989), or in the taxonomy of Oswald and colleagues (2004). Absence of the teamwork dimension in previous taxonomies might be explained by recent developments towards more group projects in the working environment, which simultaneously force universities to put more emphasize on it in their curricula.

Altogether the ETS PPI measures students' noncognitive academic potential formulated in a behavioral way. This is in line with the seminal paper of Wernimont and Campbell (1968) making the case for the measurement of past or present behavior in order to predict future behavior. Supervisor ratings of students' noncognitive academic potential might be especially valuable when predicting citizenship and counterproductive behavior, as they measure typical instead of maximal performance. Whereas standardized test results are a good predictor of maximal performance in terms of future grades, noncognitive predictors such as the ETS PPI are expected to be more appropriate to predict typical performance operationalized as citizenship and counterproductive behavior. This is in line with the idea that every cognitive and noncognitive predictor has a "natural" equivalent on the criterion side (Marcus, et al., 2007). Thus, drawing on previous research claiming that noncognitive predictors are appropriate to predict citizenship and counterproductive behavior (McHenry, Hough, Toquam, Hanson, & Ashworth, 1990; Schmitt et al., 2003) we assume:

H1: The ETS PPI is a valid predictor of students' academic potential in terms of (a) university citizenship behavior and (b) counterproductive academic behavior.

Furthermore, in order to justify the utility of supervisor ratings of students' academic potential in an applied context, incremental validity evidence above traditional predictors of study success is important (Hunsley & Meyer, 2003). Meta-analytic research shows that conscientiousness (Poropat, 2009), academic self-efficacy and achievement motivation (Robbins et al., 2004), grade goal (Richardson, Abraham, & Bond, 2012) and attitudinal constructs such as study habits (Credé & Kuncel, 2009) explain additional variance above and beyond traditional antecedents of study success. Oswald and colleagues (2004) also investigated the incremental validity of SJT and biodata measures. Both instruments provided incremental validity above and beyond traditional predictors. Notably however, both methods demonstrated stronger incremental validities when noncognitive aspects of performance such as citizenship and perseverance functioned as dependent variables. Meriac (2012) also provided evidence for the incremental validity of work ethic over traditional predictors of study success using university citizenship and academic counterproductive behavior as dependent variables. Therefore, we hypothesize the following:

H2: The ETS PPI provides incremental validity above and beyond U-GPA when university citizenship and counterproductive academic behavior function as dependent variables.

Method

Participants and Procedure

A total of 282 graduate students enrolled at a Dutch University voluntarily participated in the project. The University is located close to the German and Belgium border. The institution has a diverse student body consisting of students from those three (the Netherlands, Germany, and Belgium) and other European and non-European countries. The educational teaching system makes the university well suited to collect ratings of citizenship and counterproductive behaviors. A key element of the problem based learning teaching system is that students learn in small, self-regulated groups of 10-14 persons. Proper group functioning is dependent on team processes, such as communication, knowledge sharing, and collaboration between students. As a consequence, contextual performance aspects like citizenship and deviant behavior is critical for the functioning of these groups.

Within a four-hour testing session, the participants (1) sent ETS PPI requests to up to five former supervisors, (2) completed the GRE® revised General Test, and (3) responded to a questionnaire that included demographic information (age, gender, faculty affiliation, and undergraduate grade point average). As an incentive for participation, students were entered into a random drawing and could win one of ten iPads.

In order to send ETS PPI requests to former supervisors, students created an online ETS PPI account. They were instructed to indicate one to five former supervisors who they either knew from their Bachelor studies (university context) or from previous internships or jobs (work context). Each supervisor received an automated email inviting them to access the web-based ETS PPI system to be able to complete the evaluation.

A total of 158 of the 282 participants created an account and asked, on average, 2.7 former supervisors to fill in the ETS PPI. Thus, in total 425 ETS PPI requests were sent out and a total of 176 completed ETS PPI evaluations were received (response rate of 43%). Overall 66% of the evaluators worked at universities and 34% were employed at companies or other institutions. The final sample, for which ETS PPI evaluations were available, consisted of 115 graduate students (female, $n=61$, male, $n=54$; age = $M: 23.85$, $SD: 3.00$) enrolled at five different faculties (Business [$n=34$], Psychology [$n=58$], Law [$n=11$], Medicine [$n=7$], Arts and Social Science [$n=5$]). Participating students had received their Bachelor degree in 23 different European (Netherlands, Ireland, Germany, United Kingdom, Czech Republic, Denmark, Belgium, Bulgaria, France, Spain) and Non-European countries (South Africa, Turkey, China, Ukraine, USA, Canada, Russia, Colombia, India, Indonesia, Iran, New Zealand, Vietnam).

Approximately three months after collection of predictor data, criterion data collection took place. To that end, participants were contacted and asked to complete a questionnaire assessing their university citizenship and counterproductive academic behavior tendencies. Data thus stems from two independent sources. ETS PPI evaluations were provided by former supervisors; citizenship and counterproductive behavior was assessed by the participants themselves.

Measures

The ETS PPI. The ETS PPI consists of six dimensions: knowledge and creativity, communications skills, teamwork, resilience, planning and organization, and ethics and integrity. Each dimension is assessed using four statements (see Table 3.1). A single item overall evaluation is also provided ("Given your evaluation of the applicant of the 6 personal dimensions, please provide an overall evaluation of the applicant."). Evaluators rate the student on each statement comparing them to other students who have gone on to graduate study using a five-point scale: 1=below average, 2=average, 3=above average, 4=outstanding (top 5%), and 5=truly exceptional (top 1%). There is also an "insufficient opportunity to evaluate" field. Supervisors were instructed to rate participants in comparison to peer students within the same field of study.

Table 3.1 ETS® PPI Scales and Items**Knowledge and Creativity**

- Has a broad perspective on the field
- Is among the brightest persons I know
- Produces novel ideas
- Is intensely curious about the field

Communication Skills

- Speaks in a clear, organized, and logical manner
- Writes with precision and style
- Speaks in a way that is interesting
- Organizes writing well

Teamwork

- Supports the efforts of others
- Behaves in an open and friendly manner
- Works well in group settings
- Gives criticism/feedback to others in a helpful way

Resilience

- Accepts feedback without getting defensive
- Works well under stress
- Can overcome challenges and setbacks
- Works extremely hard

Planning and Organization

- Sets realistic goals
- Organizes work and time effectively
- Meets deadlines
- Makes plans and sticks to them

Ethics and Integrity

- Is among the most honest persons I know
- Maintains high ethical standards
- Is worthy of trust from others
- Demonstrates sincerity

Note. From ETS® Personal Potential Index. ©2009 Educational Testing Service, Princeton, NJ. ETS® Personal Potential Index is a registered trademark of ETS.

As this study is one of the first validity studies on the ETS PPI overall, and the first validity study of the ETS PPI in Europe, we did a preliminary analysis to clarify the factor structure of the ETS PPI. As a first step, we inspected zero-order intercorrelations between ETS PPI dimensions revealing strong relationships ranging from $r = .59$ to $r = .84$ ($p < .01$; see Table 2). To clarify the factor structure, we tested a one-factor model and a second-order factor model with the help of the lavaan-package (Rosseel, 2012) included in the R environment. We conducted the analysis based on all ETS PPI evaluations that were available ($N = 176$). The second-order model was defined in such a way that a higher-order ETS PPI factor accounts for the pattern of relations between six correlated lower-order ETS PPI dimensions which respectively load on the four items that were designed to measure them. Model fit was evaluated using the chi square (χ^2) test statistic, the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the standardized root mean square residual (SRMR). Results of the confirmatory factor analysis (CFA) revealed that the second-order factor model provided an acceptable model fit, $\chi^2 (248) = 448.28$; RMSEA=.07, CFI = .94, SRMR = .05, whereas a one-factor model showed an unsatisfactory fit, $\chi^2 (275) = 1023.02$; RMSEA=.12, CFI = .78, SRMR = .07. These findings imply that the six ETS PPI dimensions load on an overall factor. We accordingly operationalized the ETS PPI as an overall mean score for further analyses.

Undergraduate grade point average (U-GPA). Participants provided their undergraduate grade point average. Because students received their Bachelor degree in 23 different countries using different grading systems, they were also asked to indicate the highest and lowest possible grade underlying the grading system. Because the majority of the sample provided Dutch grades, we converted the remaining grades into the Dutch grading format (0 = lowest grade to 10 = highest grade). The grade 10.0 is hardly ever provided in the Dutch grading system (Nuffic, 2012). Therefore, the highest attainable grades in other national systems were converted into a 9.0. This conversion cannot truly equate grades assigned in different grading systems, but this rough approximation allows us to use data for students from all 23 countries in our sample.

University citizenship behavior. In order to measure university citizenship behavior, we chose a German scale (Gehring, 2006) that was developed based on research conducted in the work context (Organ, 1997; Staufienbiel & Hartz, 2000). We translated the scale into English using a translation/back-translation process and selected 14 out of 27 items that were not specifically related to the German study context, but could be easily used in the Dutch university context as well. The scale measures four central dimensions of the citizenship behavior construct. Altruism is defined as helping behavior (e.g., "I help other students having problems with understanding the subject matter."); Conscientiousness is characterized by being on time, or actively participating in group discussions (e.g., "I prepare for exams at an early stage."); Civic virtue is voluntarily behavior to support activities at the university (e.g., "I constantly inform myself about developments in academic policy."); and Sportsmanship is characterized by accepting inconvenient study conditions without complaining about it (e.g., "When the content of a non-mandatory lecture or tutorial is important for me, I go there even when the time and the location is not suitable to me."). Answers were provided on a five-point rating scale ranging from 1 = strongly disagree to 5 = strongly agree.

To examine the factor structure of the citizenship behavior construct we conducted a confirmatory factor analysis. We tested a one-factor model and a second-order model with one higher-order citizenship behavior factor accounting for the relation between four lower-order factors representing the differentiation in altruism (eight items), conscientiousness (two items), civic virtue (two items) and sportsmanship (two items). This is in line with current research (Hoffman, Blair, Meriac, & Woehr, 2007; LePine, et al, 2002) describing the citizenship behavior construct as a single factor model with underlying dimensions adding up to the latent construct. Whereas the fit of the one-factor model was unsatisfactory, $\chi^2(90) = 175.02$; RMSEA = .09, CFI = .68, SRMR = .10, the second-order factor model met the criteria for good model fit, $\chi^2(76) = 93.11$; RMSEA = .05, CFI = .94, SRMR = .07. With a value of .65, Cronbach's alpha was somewhat low. However, Cronbach's alpha has known limitations as a measure of internal consistency (Bentler, 2009; Revelle & Zinbarg, 2009). One well-known limitation is the fact that alpha is only a viable estimate of reliability when a measure is unidimensional (i.e., the correlation between the items is the same for all items). When the scale consists of some items that are more similar than others, alpha has been described as a poor estimate of the scale reliability. We therefore also estimated the reliability using an alternative measure of reliability – the greatest lower bound (glb; Woodhouse & Jackson, 1977) – that addresses this limitation of alpha. The glb has been recommended by Bentler (2009) and Sijsma (2009) and is one of several modern measures of reliability on the basis of one test administration. The glb for the university citizenship behavior scale was .83 indicating a satisfactory reliability.

Counterproductive academic behavior. In order to measure counterproductive academic behavior, we selected eight items from the established Workplace Deviance Scale (Bennett & Robinson, 2000) and adapted them to the academic context (e.g., "I take long breaks while studying." or "I make offending comments about people studying or working at my university."). Items that were too work-specific (e.g., "I discussed confidential company information with an unauthorized person." or "I falsified a receipt to get reimbursed for more money than I spent on business expenses.") and did not have a meaningful equivalent in the academic context were dropped. The scale measures two facets of counterproductive behavior: Counterproductive behavior directed towards other individuals working or studying at the university (e.g., "I make fun of other students or people working at my university.") and counterproductive behavior directed towards the university (e.g., "I am late or absent from study related meetings without permission or evident reason."). All answers were provided on a five-point rating scale ranging from 1 = never to 5 = very often.

To examine the goodness of fit of the counterproductive academic behavior construct, we also conducted a confirmatory factor analysis in the R environment. We tested a one-factor model and a second-order factor model with a higher-order general factor and two correlated first-order factors representing counterproductive behaviors directed towards the university (CWB-O; four items) and counterproductive behaviors directed towards other individuals (CWB-I; four items). A second-order factor solution was supported, $\chi^2(22) = 28.61$; RMSEA = .05, CFI = .96, SRMR = .08, whereas a one-factor solution was rejected, $\chi^2(27) = 76.10$; RMSEA = .13, CFI = .68, SRMR = .11. These findings are in line with a recent meta-analysis (Berry, Ones, & Sackett, 2007) justifying the creation of an overall counterproductive behavior construct with CWB-I and CWB-O as underlying dimensions. The internal consistency of the eight item scale was .60 (Cronbach's alpha) and .80 (glb).

Data Analysis

In order to investigate the predictive validity of the supervisor ratings of students' academic potential, we calculated zero-order correlations between all ETS PPI dimensions and the overall ETS PPI score with university citizenship and counterproductive academic behavior ratings. If more than one supervisor completed the ETS PPI, we calculated a mean score for every PPI item, and a mean score for every one of the six PPI dimensions. The mean scores for the six dimensions were averaged in one overall ETS PPI score. To compute the incremental validity of the supervisor ratings above and beyond U-GPA, we conducted a series of hierarchical regression analyses, in order to investigate the predictive validity of the ETS PPI when using (a) the overall score, (b) each dimension separately, and (c) all dimensions together. In all of these analyses, U-GPA was entered as a control variable.

Results

Table 3.2 provides an overview of means, standard deviations, and intercorrelations between study variables, and shows that the overall ETS PPI score and all separate dimensions predict university citizenship behavior after a three months lag. The same holds true for counterproductive academic behavior except for the teamwork dimension.

Table 3.2 Descriptive Statistics and Intercorrelations Between Study Variables

	<i>n</i>	<i>M</i>	<i>SD</i>	Correlations									
				1	2	3	4	5	6	7	8	9	10
1. KNOW	115	3.41	.71	—									
2. COM	115	3.35	.74	.84**	—								
3. TEAM	115	3.68	.72	.69**	.70**	—							
4. RES	115	3.55	.69	.74**	.71**	.68**	—						
5. PLAN	115	3.42	.80	.70**	.71**	.59**	.73**	—					
6. ETH	113	3.81	.75	.69**	.65**	.62**	.69**	.60**	—				
7. Overall	115	3.54	.64	.90**	.89**	.82**	.88**	.85**	.82**	—			
8. U-GPA	106	7.60	.71	.34**	.26**	.23**	.31**	.25*	.28**	.32**	—		
9. UCB	112	3.92	.42	.29**	.22*	.24*	.27**	.23*	.22*	.28**	.12 ¹	—	
10. CAB	112	3.13	.50	-.31**	-.24*	-.06	-.28**	-.36**	-.20*	-.29**	-.18 ¹	-.33**	—

Note. KNOW = Knowledge and creativity; COM = Communication skills; TEAM = Teamwork; RES = Resilience; PLAN = Planning and organization; ETH = Ethics and integrity; Overall = ETS PPI mean score aggregated across evaluators; U-GPA = Self-indicated undergraduate grade point average; UCB = University citizenship behavior; CAB = Counterproductive academic behavior. ¹ The correlations between the GRE parts and UCB and CAB were also not statistically significant.

* $p < .05$ ** $p < .01$.

Results of hierarchical regression analyses are presented in Table 3.3. In light of the high intercorrelations between ETS PPI dimensions and the results of the confirmatory factor analysis suggesting that the ETS PPI is best represented by one higher-order factor, our focus was on the overall ETS PPI score as predictor of university citizenship and counterproductive academic behavior. Results revealed that the overall score was a valid predictor of university citizenship and counterproductive academic behavior when the influence of U-GPA was controlled.

In addition to this main analysis, we conducted a series of supplementary analyses, including each ETS PPI dimension individually and including all ETS PPI dimensions jointly in a series of regression analyses. Yet, these findings may be interpreted with some caution because of the high intercorrelations between ETS PPI dimensions. Nevertheless, these supplementary analyses provide additional insights into the relationships of the individual ETS PPI subfacets with citizenship and counterproductive behavior. These analyses may be of some practical utility for universities focusing their selection procedure on specific aspects measured by the ETS PPI (e.g., communication skills) instead of simply choosing those students with the highest overall score. When entering ETS PPI dimensions individually in predicting citizenship behavior, the pattern of results was similar to the zero-order correlations depicted in Table 3.2. An exception is the ethics and integrity dimension which did not reach statistical significance when the influence of U-GPA was controlled. The same applied to counterproductive behavior. Next to the teamwork dimension that was not significantly related with counterproductive academic behavior, the ethics and integrity dimension was no longer statistically significant when controlling for U-GPA.

When all ETS PPI dimensions were included simultaneously in Step 2 of the regression analyses, knowledge and creativity was the only dimension that remained to be significant when predicting citizenship behavior. This is due to the high intercorrelations between individual dimensions. Similarly, only knowledge and creativity and planning and organization remained to be significantly negatively related to counterproductive behavior once all ETS PPI dimensions were included simultaneously. The significant positive estimate for teamwork is the result of a suppression effect as it only emerged when all ETS PPI dimensions were included simultaneously and not when they were entered individually.¹

The same pattern of results emerged when the GRE as traditional predictor of study success was included in Step 1 of the regression analyses together with U-GPA.

Table 3.3 Hierarchical Multiple Regression Analyses Predicting University Citizenship and Counterproductive Academic Behavior from U-GPA and PPI Results

Predictor	University Citizenship Behavior									
	ΔR^2	β	ΔR^2	β	ΔR^2	β	ΔR^2	β	ΔR^2	β
Step 1	.02		.02		.02		.02		.02	
U-GPA		.12		.12		.12		.12		.12
Step 2	.08**		.11***		.05*		.06**		.06*	
KNOW		—		.35***		—		—		.43*
COM		—		—		.23*		—		—
TEAM		—		—		—		.24*		—
RES		—		—		—		—		.26**
PLAN		—		—		—		—		.24*
ETH		—		—		—		—		—
Overall	.29**			—		—		—		.19
Total R ²	.09**		.12***		.07*		.08**		.07*	.14*
Counterproductive Academic Behavior										
Step 1	.03		.03		.03		.03		.03	
U-GPA		-.18		-.18		-.18		-.18		-.18
Step 2	.08**		.10***		.05*		.08**		.04	
KNOW		—		-.33***		—		—		-.42*
COM		—		—		-.23*		—		—
TEAM		—		—		—		—		.15
RES		—		—		-.05		—		—
PLAN		—		—		—		-.29**		—
ETH		—		—		—		—		-.38***
Overall	.30**		.13***		.09*		.11**		.17***	.20
Total R ²	.11**		.13***		.09*		.04		.07	.26***

Note. $N = 106$. U-GPA = Self-indicated undergraduate grade point average; KNOW = Knowledge and creativity; COM = Communication skills; TEAM = Teamwork; RES = Resilience; PLAN = Planning and organization; ETH = Ethics and integrity; Overall = ETS PPI mean score aggregated across evaluators * $p < .05$ ** $p < .01$ *** $p < .001$.

Discussion

Results of the present study support the usefulness of supervisor ratings of students' academic potential as assessed in the ETS PPI to predict university citizenship and counterproductive academic behavior. This is in line with findings from the work context that supervisor' ratings are reliable predictors of job performance (Viswesvaran, Ones & Schmidt, 1996). Furthermore, it broadens existing knowledge obtained in the academic context showing that evaluations provided by former supervisors are valid predictors of future grades (Aamodt, Bryan, & Whitcomb, 1993) and peer-ratings of academic performance (Stanton, Burstein, Kobos, & Loucks, 1979). U-GPA, however, did not explain a significant amount of variance within the outcome measures. Although meta-analyses have confirmed the predictive validity of U-GPA as indicator of study success (Kuncel, Credé, & Thomas, 2007; Kuncel, Hezlett, & Ones, 2001), it appears not to be a valid predictor of university citizenship and counterproductive academic behavior. Thus, results provide evidence for the incremental validity of the ETS PPI, as it predicts university citizenship and counterproductive academic behavior above and beyond U-GPA.

Supervisor ratings of students' academic potential as measured by using the ETS PPI will likely function as supplemental information rather than as a substitute for cognitive instruments. Thus, the incremental validity evidence we found in the present study is an important premise to justify their utility in an applied context. Combining multiple predictors has been described as a viable method to find a balance between validity and diversity of selection instruments (Sackett & Lievens, 2008). This idea is based on the assumption that noncognitive correlates, such as interests or personality, produce less subgroup differences compared to standardized tests that target cognitive skills (Hough, Oswald, & Ployhart, 2001; Sackett, Schmitt, Ellingson, & Kabin, 2001). Thus, in case admission committees aim to base their selection decisions on a broader conceptualization of academic performance in order to increase the diversity of the student body (Ployhart & Holtz, 2008; Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997), the use of standardized supervisor ratings seems to be an appropriate approach.

Previous studies investigating the incremental validity of noncognitive predictors primarily used students' grade point average as the dependent variable (Chapman & Hayslip, 2005; Farsides & Woodfields, 2003; Gore, 2006). Doing so, they found mixed results for the incremental validity of noncognitive instruments. The findings of the present study, however, underscore the need to extend the criterion space when investigating the incremental validity of noncognitive predictors to include aspects of academic performance that go beyond task performance. This is in line with Enright and Gitomer (1989) who argued that it is not sufficient to evaluate assessments in terms of their ability to predict future grades. Accordingly, we conclude that future validity studies should not only extend the predictor, but also the criterion domain, in order to prevent underestimation of incremental validity evidence.

The validity evidence of the ETS PPI and the resulting possibility for its use in graduate selection has several practical implications. First of all, our study findings suggest that the ETS PPI is a valid and useful instrument for student selection. The ETS PPI may therefore be of great practical value to organizations as it is also efficient and time-effective for supervisors; supervisors only have to fill in one evaluation per student independent of the number of universities for which an applicant wants to apply. The student can then designate which universities the ETS PPI report is to be sent to and staff members will not have to write individual letters of recommendation or complete various forms required by the respective universities. Second, completing the ETS PPI is quicker than writing letters of recommendation, since the supervisor is responding to the same set of 24 statements across all students. Third, the standardization makes it easier for universities to compare students applying from different universities and even different countries. Students' mobility has considerably increased within Europe and applying students come from various countries. Thus, the standardized wording of the ETS PPI avoids possible misinterpretations due to cultural and language differences (Liu, et al., 2009).

Taken together, findings of the present study are in line with Meriac (2012) who provided evidence for the positive relation between work ethic and university citizenship behavior, and the negative relation with counterproductive academic behavior. Although surprisingly the ethics and integrity dimension was not a valid predictor of counterproductive behavior when the influence of U-GPA was held stable, we provided evidence that overall supervisor ratings of students' academic potential as measured by the ETS PPI predict citizenship and counterproductive behavior in the academic context. Furthermore, we extended previous research (Meriac, 2012; Oswald et al., 2004; Zettler, 2011) by predicting student citizenship and counterproductive academic behavior using supervisor ratings instead of self-ratings.

Limitations and Future Directions

The present study is not without limitations; in the remainder, we critically discuss these limitations and identify potential areas of future research. First, the current study is based on a relatively small sample size. Future research might investigate the validity of the ETS PPI in a large-scale study. This would also allow comparisons to be made between different subsamples (e.g., female vs. male; European vs. Non-European) in order to evaluate the extent to which including the ETS PPI in a selection procedure reduces subgroup differences. Furthermore, the time between predictor and criterion data collection might be extended in future research, to investigate whether this has an influence on the validity of the ETS PPI.

Second, another shortcoming is that the present study was conducted in a low-stakes testing situation because students had already been admitted to their study program. This has both advantages and disadvantages: A disadvantage is that being aware that the student has already been accepted into a master's program may have caused biased supervisor ratings. It may therefore be valuable to replicate the present findings in a high-stakes testing situation. It would be especially interesting to compare findings from a high-stakes situation with the present study findings to investigate whether supervisors evaluate students systematically higher if students need the evaluation for an actual application. On the other hand, our procedure also had advantages as range restriction in the predictor is less of a concern. Our sample of students had not gone through a specific selection procedure before they could register for their master's program. Most faculties at the university used in the study required only verification of students' Bachelor degrees and a proof of their English proficiency level. As the ETS PPI was not part of this procedure we did not have range restriction in the predictor. In contrast, graduate programs in other countries are much more selective only admitting a small subset of applicants. Thus, predictive validity studies conducted in more competitive educational systems might lead to lower levels of predictive validity due to range restriction underlying the need for further validity studies in other educational systems.

Third, one may consider the use of self-ratings of university citizenship and counterproductive academic behavior a shortcoming of the present study. However, research on citizenship and counterproductive behavior frequently relies on self-ratings (Fox, Spector, Goh, & Bruursema, 2007), because behaviors such as cheating or lying are mostly covert, and therefore remain unrecognized by others, making it difficult to assess counterproductive behavior with other-ratings. Recent meta-analyses comparing self- and other-ratings of citizenship and counterproductive behavior have also provided support for the utility of self-ratings (Berry, Carpenter, & Barratt, 2012; Carpenter, Berry, & Houston, 2014). Furthermore, the present study relied on multi-source data (former supervisor ratings for the ETS PPI and self-ratings for citizenship and counterproductive behavior) thereby reducing common method bias. In line with a study conducted in the work context (Connelly & Hülshager, 2012), future research might also consider the possibility to predict noncognitive aspects of performance by using other-ratings from outside the university context (e.g., family and friends).

Fourth and finally, a general shortcoming that is inherent in selection procedures such as letters of recommendations or the ETS PPI is the leniency bias (Aamodt, 1999). The likelihood that former supervisors provide favorable instead of honest ratings is high, as applicants themselves choose which former supervisors provide the information. In order to reduce this problem the answer scale of the ETS PPI was designed in such a way that it was imbalanced (Kyllonen, 2008). Assuming that supervisors provide their ratings on the positive side of the scale, this side was provided with a larger number of response choices: Three answer possibilities are above average and only one is below. Although empirical findings (Kyllonen, 2008) and the results of the present study show that supervisors are still prone to provide positive ratings (Kyllonen: $M = 3-4$; present study: $M = 3.57$), the ETS PPI allows for a more precise differentiation between highly-ranked applicants. However, future research is needed in order to further investigate ways to reduce the influence of leniency biases and increase fairness in student selection.

Chapter 4

Be Aware to Be on the Square: Mindfulness and Counterproductive Academic Behavior

Published as:

Schwager, I. T. L., Hülshager U. R., & Lang, J. W. B. (in press). Be Aware to Be on the Square: Mindfulness and Counterproductive Academic Behavior. *Personality and Individual Differences*.

Abstract

The aim of the present study was to investigate the role of dispositional mindfulness – the capacity to be nonjudgmentally aware of the present moment (Brown & Ryan, 2003) – in counterproductive academic behavior. Apart from investigating the direct relationship between mindfulness and counterproductive behavior, we tested the moderating role of established personality dimensions (conscientiousness and honesty-humility) in the relationship between mindfulness and counterproductive academic behavior. Two hundred eighty-two graduate students completed a trait mindfulness measure and a personality inventory based on the HEXACO model, followed by self-ratings of counterproductive academic behavior after a three-month time lag. Hierarchical regression analyses revealed that mindfulness, conscientiousness, and honesty-humility were negatively related to counterproductive academic behavior. As hypothesized, conscientiousness and honesty-humility moderated the relationship between mindfulness and counterproductive behavior, such that the mindfulness-counterproductive behavior relationship was stronger for students low on conscientiousness and honesty-humility. These findings add to previous findings on the positive effects of mindfulness for students by demonstrating that it also benefits professional academic behavior.

Keywords: mindfulness, counterproductive academic behavior, HEXACO model of personality

Plagiarism, cheating, absenteeism, substance abuse, stealing, or procrastination – the list of counterproductive academic behaviors is long. Counterproductive behaviors are negatively related to students' academic achievements in terms of grades (Credé & Niehorster, 2009) and hamper other group members' performance (Koppenhaver, 2006). Cheating, for example, prevents students from understanding the content matter, whereas absenteeism reduces the possibilities to learn from each other not only for the absent student, but also for the rest of the work group (Koppenhaver, 2006). Likewise, universities' well-functioning and reputation may suffer from graduate students engaging in counterproductive behavior underlining the need to identify mechanisms through which academic misbehavior can be reduced.

Previous research investigating antecedents of counterproductive academic behavior has focused on the role of personality traits as specified in the Big Five and HEXACO model. Specifically, conscientiousness and honesty-humility were established as valid predictors of counterproductive academic behavior (de Vries, de Vries, & Born, 2011; Marcus, Lee, & Ashton, 2007). In the present study, we focus on trait mindfulness, an individual difference variable that has, to our knowledge, not yet been investigated in relation to counterproductive (academic) behavior but which has the potential to provide additional insights into the antecedents and processes involved in counterproductive behavior. Mindfulness describes a state of consciousness in which individuals bring awareness to what is occurring in the present moment with a nonjudgemental attitude (Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006; Brown & Ryan, 2003). Although conceptually mindfulness is a state, researchers agree that there are rather stable trait-like between-person differences in the extent and frequency with which individuals experience mindful states (Brown & Ryan, 2003; Glomb, Duffy, Bono, & Yang, 2011). Although mindfulness displays meaningful relationships with traditional Big Five personality traits (Giluk, 2009; Thompson & Waltz, 2007), it also differs from established personality traits and is worthy of investigation in its own right (Brown & Ryan, 2003; Giluk, 2009).

Goal of the present study is to link mindfulness to counterproductive academic behavior. Doing so, we will build upon affective events theory (AET; Weiss & Cropanzano, 1996). Specifically, we argue that mindfulness displays an overall negative relationship with counterproductive academic behavior. In addition, we posit that the personality characteristics of conscientiousness and honesty-humility shape this overall relationship such that the relationship is stronger when conscientiousness and honesty-humility are low.

Mindfulness and Counterproductive Academic Behavior

The link between mindfulness and counterproductive behavior can be understood against the backdrop of AET (Weiss & Cropanzano, 1996; Matta, Erol-Korkmaz, Johnson, Bicaksiz, 2014) originating in the industrial and organizational (I/O) psychology literature. Accordingly, work events trigger affective experiences, which, in turn shape employees work behavior, including counterproductive behavior. Transferring this idea to the context of academia, AET suggests that counterproductive academic behavior is driven by students' affective states such as anger, anxiety or fear.

Linking mindfulness to extant work on counterproductive behavior suggests that mindfulness is negatively related to counterproductive student behavior by helping students to better regulate affective states. Previous research conducted in the work context indicates that trait positive and trait negative affect are related to counterproductive behavior (Dalal, 2005). High positive affect is associated with high levels of energy, full concentration, and pleasurable engagement, whereas high negative affect is associated with subjective distress and unpleasurable engagement (Watson, Clark, & Tellegen, 1988). Thus, it might seem intuitively appealing that people with high positive affect engage less in counterproductive behavior, whereas people with high negative affect show counterproductive tendencies more frequently.

Mindfulness at the same time is negatively related to negative affect (Brown & Ryan, 2003; Giluk, 2009; Schutte & Malouff, 2011) and facilitates adaptive forms of emotion regulation (Baer, et al., 2006; Brown & Ryan, 2003; Hülshager, Alberts, Feinholdt, & Lang, 2013). A study investigating the neurocognitive underpinnings of trait mindfulness demonstrated that more mindful individuals showed increased activity in brain regions associated with the regulation of negative affect when they had to label emotionally threatening stimuli (Creswell, Way, Eisenberger, & Lieberman, 2007). The authors conclude that the beneficial effects of mindfulness might be explained through the ability to label one's negative emotions which in turn reduces the intensity of the emotion, and finally the likelihood to engage in automatic maladaptive behavior in response to it. This is in line with empirical findings indicating that mindfulness is negatively related to the use of avoidant-oriented coping strategies (Weinstein, Brown, & Ryan, 2009). Thus, we hypothesize the following:

H 1: Mindfulness is negatively related to counterproductive academic behavior.

Considering Conscientiousness and Honesty-Humility as Boundary Conditions

Considering the boundary conditions under which mindfulness might lessen the occurrence of counterproductive academic behavior, the role of established personality traits as assessed by the Big Five and HEXACO model needs to be taken into account. In the I/O psychology literature, evidence has accumulated that honesty-humility is the strongest predictor of workplace delinquency (Lee, Ashton, & de Vries, 2005a; Lee, Ashton, & Shin, 2005b; Oh, Lee, Ashton, & de Vries, 2011). Conscientiousness has also been shown to be negatively related with workplace deviance (Berry, Ones, & Sackett, 2007; Lee et al., 2005b; Salgado, 2002). Similarly, research conducted in the academic context provides evidence that honesty-humility and conscientiousness are negatively correlated with counterproductive academic behavior (de Vries, et al., 2011; Marcus et al., 2007). Students characterized by high levels of honesty-humility can be described as honest, sincere and fair (Lee & Ashton, 2004). Conscientious students are considered to be well organized, disciplined, precise and self-controlled (Lee & Ashton, 2004). Whereas the negative relation between honesty-humility and counterproductive behavior appears to be straightforward, conscientious students might engage less in counterproductive academic behavior as they are more likely to exert themselves for reaching their task-related goals (e.g., studying to pass an exam) instead of choosing a counterproductive alternative (e.g., cheating to pass an exam; de Vries, et al., 2011). Taken together, this line of research suggests that students with a high dispositional tendency towards conscientiousness and honesty-humility have a low overall tendency to engage in counterproductive academic behavior. For these highly conscientious and honest students, a ceiling effect may occur leaving little room for mindfulness to exert a positive influence. Thus, students with high honesty-humility and conscientiousness are expected to display low counterproductive academic behavior irrespective of their level of trait mindfulness. In contrast, the positive potential of mindfulness may fully unfold in the case of students low on conscientiousness and honesty-humility. Due to their personality disposition, they are inclined to engage in counterproductive academic behavior. This inclination may, in turn, be hampered by mindfulness such that the relationship between mindfulness and counterproductive academic behavior may be strong when students are at risk of displaying counterproductive behavior or due to low levels of conscientiousness and honesty-humility. Therefore, we hypothesize the following:

H 2a: Conscientiousness moderates the relationship between mindfulness and counterproductive academic behavior, such that the negative relationship is stronger when conscientiousness is low.

H 2b: Honesty-humility moderates the relationship between mindfulness and counterproductive academic behavior, such that the negative relationship is stronger when honesty-humility is low.

Method

Participants and Procedure

The overall sample consisted of 281 master students (Age: $M = 23.55$; $SD = 2.36$; female = 154; male = 127) enrolled at five different faculties (Business, Psychology, Law, Health, Medicine and Life Science, and Arts and Social Science) of a Dutch University. The data used in this study was collected as part of a larger 3.5-year research project on cognitive and non-cognitive predictors of study success. Other publications that resulted from this project are Schwager, Hülshager, Bridgeman, & Lang (2015) on the predictive validity of the General Record Examination (GRE) for student task performance and Schwager, Hülshager, Lang, Klieger, Bridgeman, & Wendler (2014) on the predictive validity of the Personal Potential Index (PPI), a non-cognitive student selection instrument, for citizenship and counterproductive academic behavior. Theoretical ideas and analyses reported in the present study do not overlap with either of these publications.

Students participated voluntarily and provided their informed consent. Those students being interested were entered into a lottery and could win one of ten iPads. They had just started their master program when filling in the personality and mindfulness questionnaire. Counterproductive academic behavior was assessed approx. three months later, when students had finished the first teaching term comprising lectures, tutorial group meetings and at least one exam. This is considered to be advantageous as a time lag between the investigation of predictor and criterion variables helps to reduce the problem of common method bias (Podsakoff, MacKenzie, Lee & Podsakoff, 2003). In total, two-hundred sixty-one students filled in the questionnaire at time 2.

Measures

Mindfulness. Trait mindfulness was assessed with the Mindful Attention Awareness Scale (MAAS; Brown & Ryan, 2003). All 15 items were reversed coded in such a way that they measure how often participants are in mindless states (e.g., "I find myself doing things without paying attention."). Accordingly, all items were recoded so that high values corresponded with high dispositional mindfulness before an average score was calculated. The rating scale ranged from 1 = almost always to 6 = almost never. Cronbach's alpha measuring the internal consistency across the 15 items measuring trait mindfulness was .80.

Personality. Personality traits were measured by using the HEXACO-60 (Ashton & Lee, 2009). The HEXACO-60 assesses a set of six broad personality dimensions: honesty-humility, emotionality, extraversion, agreeableness, conscientiousness and openness to experience. Taking into account all six HEXACO dimensions allowed us to follow recent calls to control for stable personality characteristics when investigating mindfulness in relation to outcome criteria. This was deemed important in light of the conceptual overlap between mindfulness and some Big Five personality traits (Giluk, 2009). Items that were negatively formulated were recoded before a mean score was calculated for each of the six dimensions. Each dimension was made up of ten items and all items were rated on a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree). The internal consistency (Cronbach's alpha) across the items measuring the six personality dimensions ranged from .66 for agreeableness to .78 for conscientiousness and extraversion.

Counterproductive Academic Behavior. Counterproductive academic behavior was investigated with eight items, extracted from the Workplace Deviance Scale (Bennett & Robinson, 2000). Items were adapted to the academic context ("I am late or absent from study related meetings without permission or evident reason."; "I spend a lot of time with daydreaming or losing myself in thoughts instead of studying."; "I make offending comments about people studying or working at my university."; "I take long breaks while studying."). Answers were provided on a five-point rating scale (1 = never to 5 = very often).

To examine the goodness of fit of the counterproductive academic behavior construct, we conducted a confirmatory factor analysis. We tested a second-order factor model with a higher-order general factor and two correlated first-order factors representing counterproductive behaviors directed towards the university (CWB-O; four items) and counterproductive behaviors directed towards other individuals (CWB-I; four items) against a first-order model with two correlated factors representing CWB-O and CWB-I. Both models had an acceptable model fit: χ^2 (22) = 52.739; RMSEA=.073, CFI = .926, SRMR = .070 (second-order model); χ^2 (25) = 52.739; RMSEA=.065, CFI = .934, SRMR = .070 (first-order model). In line with previous research justifying the creation of an overall counterproductive behavior construct with CWB-I and CWB-O as underlying dimensions (Berry, et al., 2007; Marcus, Taylor, Hastings, Sturm, & Weigelt, 2013), we aggregated all items in one overall score. The internal consistency across the eight items, as assessed by Cronbach's alpha, was .70.

Results

Table 1 provides an overview of the zero-order correlations between the study variables. Dispositional mindfulness was negatively related to counterproductive academic behavior ($r = -.28, p < .01$) after a three months' time lag. With respect to personality, conscientiousness ($r = -.45, p < .01$), agreeableness ($r = -.13, p < .05$), emotionality ($r = -.17, p < .01$) and honesty-humility ($r = -.29, p < .01$) were significantly correlated with subsequent counterproductive academic behavior. Extraversion and openness were not significantly correlated with counterproductive behavior.

Results concerning Hypothesis 1 revealed that mindfulness ($\beta = -.13, p < .05$) was negatively related with counterproductive academic behavior when the influence of the six HEXACO personality dimensions was held stable. Furthermore, conscientiousness ($\beta = -.40, p < .01$), honesty-humility ($\beta = -.17, p < .01$), and emotionality ($\beta = -.13, p < .05$) shared a significant amount of variance with counterproductive academic behavior (see Table 2).

In order to test Hypotheses 2a and 2b, we calculated the interaction terms between (1) mindfulness and conscientiousness, and (2) mindfulness and honesty-humility. To reduce the problem of multicollinearity, we standardized the predictor variables before the interaction terms were calculated (Marquardt, 1980). The standardized interaction terms were first entered separately into the second step of the regression analysis (see Table 2). In addition, we report an analysis in which they are entered jointly. When included separately, conscientiousness ($\beta = .12, p < .05$) and honesty-humility ($\beta = .13, p < .05$) significantly moderated the relationship between mindfulness and counterproductive behavior. When both interaction terms were included jointly in the second step of the regression analysis, only the interaction effect between mindfulness and honesty-humility was significant ($\beta = .11, p < .05$).

Table 4.1 Descriptive Statistic, Reliabilities and Intercorrelations Between Study Variables

	<i>n</i>	<i>M</i>	<i>SD</i>	<i>α</i>	Correlations							
					1	2	3	4	5	6	7	8
1. MIND	281	4.30	0.61	.80	—							
2. CO	281	3.60	0.58	.78	.35**	—						
3. AG	281	3.10	0.52	.66	.20**	.01	—					
4. EM	281	3.07	0.59	.75	-.11	.08	-.01	—				
5. HO	281	3.37	0.57	.71	.14*	.15*	.31**	.15*	—			
6. EX	281	3.60	0.55	.78	.24**	.11	.08	-.14*	-.11	—		
7. OP	281	3.56	0.62	.75	.04	-.01	.08	-.03	.17**	.17**	—	
8. CAB	261	1.85	0.48	.70	-.28**	-.47**	-.14*	-.17**	-.29**	.04	-.04	—

Note. MIND = Mindfulness; CO = Conscientiousness; AG = Agreeableness; EM = Emotionality; HO = Honesty-Humility; EX = Extraversion; OP = Openness to Experience; CAB = Counterproductive academic behavior.
* $p < .05$ ** $p < .01$.

Table 2
Moderated Regression Analysis Predicting Counterproductive Academic Behavior from Mindfulness and Personality

Predictor	Model 1		Model 2		Model 3		Model 4	
	ΔR^2	β	ΔR^2	β	ΔR^2	β	ΔR^2	β
Step 1	.31***		.31***		.31***		.31***	
Mindfulness		-.13*		-.13*		-.11		-.12*
Conscientiousness		-.40***		-.38***		-.38***		-.37***
Agreeableness		-.08		-.08		-.07		-.07
Emotionality		-.13*		-.13*		-.13*		-.13*
Honesty-humility		-.17**		-.15**		-.16**		-.15**
Extraversion		.11		.11		.10		.10
Openness		-.04		-.04		-.03		-.03
Step 2			.01*		.02*		.02*	
Mindfulness x conscientiousness				.12*				.09
Mindfulness x honesty-humility						.13*		.11*
Total R^2	.31***		.32*		.33*		.33*	

Note. $n = 261$.
* $p < .05$ ** $p < .01$ *** $p < .001$.

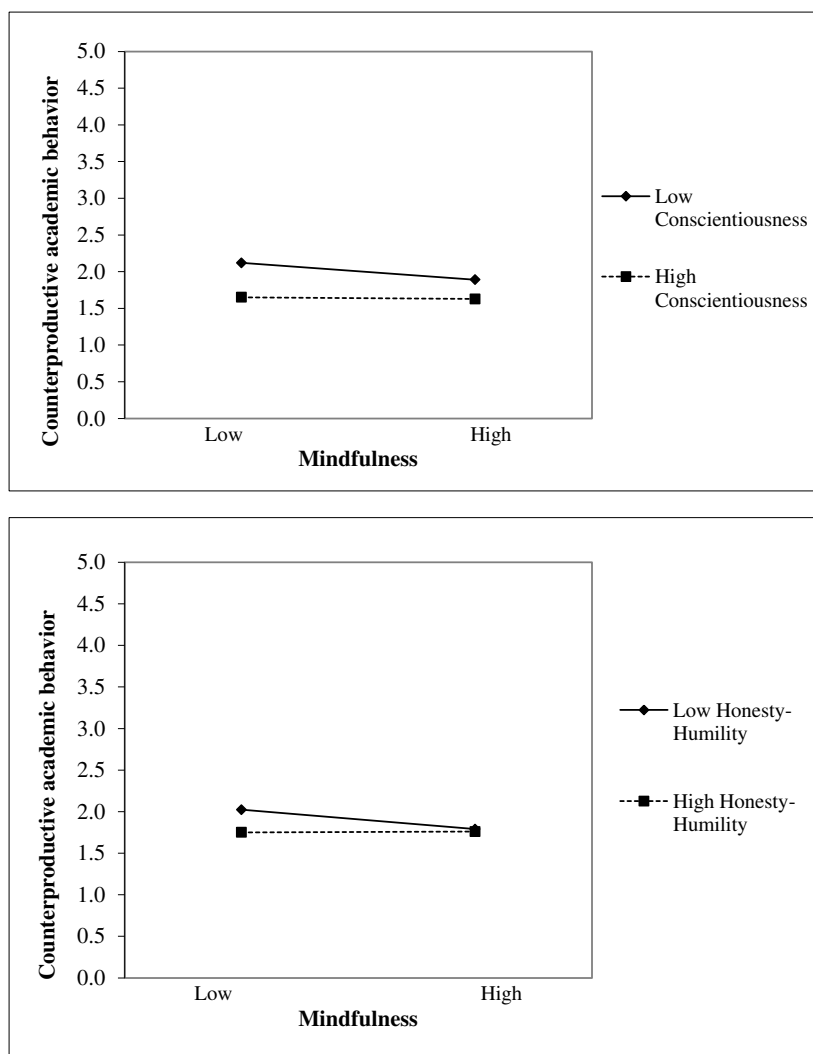


Figure 4.1 Interactions of mindfulness with conscientiousness and honesty-humility on counterproductive academic behavior (CAB)

A graph of the significant interactions is provided in Figure 1. Results of simple slope analyses (Aiken & West, 1991) revealed that the relationship between mindfulness and counterproductive academic behavior was significant when conscientiousness ($B = -.11$, $p < .05$) and honesty-humility ($B = -.12$, $p < .01$) were low (-1 SD). In contrast, the relationships were not significant, when conscientiousness ($B = .01$, n.s.) and honesty-humility ($B = .01$, n.s.) were high ($+1$ SD).

Discussion

The primary aim of the present study was to investigate the relation between mindfulness and counterproductive academic behavior. Results indicate that overall higher trait mindfulness was associated with less counterproductive academic behavior measured after a three months time lag. Although we did not explicitly test the mechanisms underlying this relationship, findings are in line with predictions made by AET (Weiss & Cropanzano, 1996). Accordingly, students high vs. low on mindfulness may differ in their perception and appraisal of events taking place in the academic context. This may trigger different affective reactions which may ultimately lead to differences in counterproductive behaviors: More mindful students may perceive demanding situations as less stressful and threatening (Weinstein, et al., 2009) and this may help them focus on the fulfillment of the task itself ("What can I do right now in order to pass an exam?") rather than engaging in counterproductive behaviors. In contrast, students low on mindfulness are occupied by worry-related thoughts about the past or the future (e.g., "What happens if I don't pass the exam?"). These thoughts take away important resources, they trigger negative affective reactions which may influence their behavior towards other students and lead to counterproductive behaviors such as cheating, gossiping, withholding information, etc..

With respect to traditional personality traits, conscientiousness and honesty-humility were valid predictors of future counterproductive behavior. The negative link between honesty-humility and counterproductive behavior is in accordance with the findings of a recent study (Meriac, 2012) linking work ethics to counterproductive academic behavior. The negative relation between emotionality and counterproductive behavior is also in line with previous findings from the work context (Lee et al., 2005a). In contrast to neuroticism which is characterized by anxiety, fear, impulsivity and worry in the Big Five taxonomy (Costa & McCrae, 1992), emotionality measured by the HEXACO focuses on empathy and the need for emotional support from others (Ashton & Lee, 2007). This might explain why emotionality is negatively related with counterproductive behavior in contrast to neuroticism which is positively related with counterproductive behavior in previous research (Berry et al., 2007).

Although we found an overall negative relationship between mindfulness and counterproductive behavior, the interaction of mindfulness with conscientiousness and honesty-humility indicated that this relationship was predominantly driven by individuals low on conscientiousness and honesty-humility. Thus results suggest that whereas students with high levels of conscientiousness and honesty-humility have little inclination to show counterproductive behavior regardless of their level of mindfulness, students with a strong tendency towards unreliability and dishonesty profit from high levels of trait mindfulness.

Practical Implications

Finding a relation between mindfulness and counterproductive academic behavior is promising from universities' perspective. Compared to the relatively stable personality traits defined in the Big Five and HEXACO model, trait mindfulness can be enhanced through training (Carmody & Baer, 2008; Shapiro, Brown, & Biegel, 2007). This is in line with empirical research indicating that mindfulness-based interventions focusing on stress-reduction have the potential to increase trait mindfulness levels and thereby decrease stress symptoms of university students (Regehr, Glancy, & Pitts, 2013). Furthermore, previous research has revealed that trait mindfulness is positively related to performance in terms of grades, test results, information retention and working memory capacity (Mrazek, Franklin, Philipps, Baird, & Schooler, 2013; Ramsburg & Youmans, 2013). Thus, universities might consider providing mindfulness trainings in order to reduce students' tendencies to engage in counterproductive behavior while simultaneously making a positive contribution to their coping potential and task performance.

Limitations and Future Research Directions

The first limitation of the present study is that all ratings stem from the same data source (common method bias). Yet, while the validity of observer ratings of traditional personality traits is well-established (Connelly & Ones, 2010) this is not the case for mindfulness and counterproductive behavior for which other-ratings might be less accurate than self-reports. Whereas public counterproductive behaviors such as absenteeism can be observed by peers or supervisors, hidden behaviors such as cheating or daydreaming remain unrecognized by others. The same holds true for mindfulness which refers to internal processes and experiences which are not readily observable by others.

A second potential limitation is that the relation between mindfulness and counterproductive behavior found in the present study is only cross-sectional in nature. Although students were asked to provide information about their counterproductive tendencies three months after trait mindfulness and personality traits were assessed, this set-up does not allow drawing causal conclusions. A more rigorous test of directions of effects would involve a full two-wave panel design (Zapf, Dormann, & Frese, 1996) or an experimental design making use of a mindfulness intervention (as e.g., in Hülshager et al., 2013).

Notably, the magnitudes of the interaction effects were relatively small. However, interaction effects are especially difficult to detect due to statistical issues such as range restriction in the predictor (Aguinis, 1995), and reduced reliability of the product term (Dunlap & Kemery, 1988). Furthermore, the distribution of the dependent variable was skewed towards low ratings which may further limit the probability to detect interaction effects (Hamilton, 1992). This might be due to the fact that master students are usually well-motivated and therefore engage less often in counterproductive behavior as for example undergraduate students. Future research may thus focus on replicating our findings using different types of samples (e.g., undergraduate or college students).

The present study confirmed conscientiousness and honesty-humility as moderators of the mindfulness-counterproductive behavior relationship. This finding does not preclude the possibility that there are other important moderators of this relationship. While we focused on moderators that are related to the individual, situational conditions are also likely to moderate the mindfulness-counterproductive behavior relationship. For instance, mindfulness may be especially important when study conditions are stressful or perceived as unfair and consequently associated with high levels of negative affective experiences. In contrast, the positive potential of mindfulness may unfold to a lesser extent when study conditions are perceived as optimal and trigger few negative affective experiences.

Chapter 5

General Discussion

The world of work has changed in the 21st century. Whereas people from previous generations have often worked in one job or even for one employer for their entire lifetime, students entering the job market of today are likely to work for different employers and even in different jobs. Thus, the importance of noncognitive qualities that promote individuals' sustainable employability has considerably increased. Consequently, universities aiming to prepare students for the job market try to adapt to this development by not solely transferring knowledge, but also fostering students' noncognitive skills. This is in line with a recent trend in the literature to conceptualize academic performance in broader terms encompassing not only task performance (e.g., GPA), but also citizenship (e.g., helping other students) and counterproductive behavior (e.g., cheating). Based on empirical data collected at a Central European university, this dissertation contributes to this stream of literature by providing meaningful insights into how to obtain master students that are suitable with respect to all three performance aspects.

Although the need to take students' noncognitive performance into account has been raised increasingly often in the last decade, the vast amount of literature investigates how to predict students' cognitive performance. This is in line with a review conducted by Sackett, Bornemann and Connelly (2008) in which they integrated empirical findings that primarily refer to cognitive loaded tests while simultaneously covering the question of how to predict a broadened criterion domain. In doing so, they summarized eight main assertions that are often expressed with respect to high-stakes testing in the domain of higher education and respond to the criticism by integrating findings from large scale studies and meta-analyses. Based on the reported findings, they concluded that none of the assertions has been supported. In the following, I will critically discuss in how far the empirical findings obtained in this dissertation further help to refute some of these assertions (see Table 5.1).

Table 5.1 Criticism Raised in the Field of High-Stakes Testing

Assertion	Description
1. Tests predict badly, if at all	Correlations between standardized tests and short term criteria such as GPA typically vary between .25-.35. Thus the accounted variance is considered to be small.
2. Tests do not measure all important determinants of all important criteria	
a) Tests are not predictive of earnings	a) Although income is a criterion of real interest, standardized tests are considered not to be predictive of this outcome measure.
b) Some relevant criteria are not predicted well by tests of developed abilities	b) Standardized tests do not predict noncognitive performance aspects such as citizenship and counterproductive behavior.
c) Important determinants of traditional criteria are missed	c) Noncognitive predictors explain additional variance above and beyond cognitive loaded tests when predicting ability related outcome variables such as GPA.
3. The appearance of validity is due to SES	High SES has an influence on tests scores and grades. When the influence of SES is controlled for the relationship disappears.

Retrieved from Sackett et al., (2008)

The first assertion “tests predict badly, if at all” (Sackett et al., 2008; p. 215) refers to the magnitude of the correlation coefficients between standardized tests and commonly used outcome measures (see Table 5.1). In the present study the uncorrected zero-order correlations between GRE results and G-GPA ranged between .17 (GRE-Q and G-GPA), .21 (GRE-V and G-GPA) and .31 (GRE-AW and G-GPA). Assuming that unreliable criterion measures lead to an underestimation of validity evidence, I corrected for measurement error in the criterion (i.e., G-GPA). As a result, the correlations increased by .06 to .23 (GRE-Q and G-GPA), by .07 to .28 (GRE-V and G-GPA) and by .10 to .41 (GRE-AW and G-GPA) respectively. Thus taking into account the corrected correlations, GRE-Q explained 5%, GRE-V 8% and GRE-AW 10% of the variance in students’ academic performance measured in terms of G-GPA. Taken together the three GRE parts accounted for 15% in the criterion. These findings are comparable to findings from other predictive validity studies (Sackett et al., 2008). However, in line with the aforementioned assertion, some researchers consider these effects to be limited and question the usefulness of the validities (Brown, 2007; Sherley, 2007). In response to that, it has been argued that predictive validity coefficients are often misinterpreted when the high practical significance of small values is underestimated (Bridgeman, Burton, & Cline, 2009). Thus, Bridgeman and colleagues (2009) analyzed whether students’ success rates operationalized in terms of grades differ depending on their GRE scores. In doing so they divided the sample into three quartiles based on their study success: One representing the top quarter of students (most successful students), the middle quarter (average students) and the bottom quarter (least successful students). Results revealed that 32% of the students scoring low on the GRE also showed low success rates during their studies. In order to differentiate between successful students, especially high GRE results provided meaningful information. Whereas students with average and low GRE scores were almost equally likely to obtain good grades, 39% of the students with high GRE results were successful during their studies. This alternative method of presenting the predictive power of the GRE together with the findings of previous predictive validity studies (Kuncel, Hezlett, & Ones, 2001; Kuncel, Wee, Serafin, & Hezlett, 2010) and the findings presented in Chapter 2 underscore the usefulness of the GRE as a selection instrument that provides meaningful information when predicting students’ GPA.

Under the second assertion “tests do not measure all important determinants of all important criteria”, Sackett and colleagues (2008; p. 215) summarized different points of critique (see Table 5.1). The first point stating that standardized tests do not predict earnings is based on the idea that income is the criterion of real interest. Although the focus of this dissertation was on the prediction of short-term criteria such as G-GPA, information about participants’ salary after they have entered the job market has been collected. Results of the hierarchical regression analyses controlling the influence of faculty affiliation revealed that GRE-AW is a valid predictor of income ($\beta = .21$, $p < .05$). Nonetheless, Sackett and colleagues argue that primarily other factors than students’ aptitudes explain how much money they earn (e.g., choice of occupation).

The same might hold true for degree attainment. Although it has been assumed that the GRE predicts whether students manage to obtain their master degree in the predefined amount of time, results presented in this dissertation reveal that the relationship between GRE and degree attainment was not significant. Thus, other factors than students' cognitive abilities need to be considered when predicting criteria such as income or degree attainment that are not primarily ability-related.

Another point of critique blames standardized tests for not being predictive of non-cognitive performance aspects such as citizenship and counterproductive behavior (see Table 5.1). In order to investigate whether the PPI has practical utility above and beyond traditional predictors of study success when predicting citizenship and counterproductive behavior, the influence of U-GPA and GRE has been controlled for in Chapter 3. Results indeed revealed that both U-GPA and GRE did not predict citizenship and counterproductive behavior in the academic context. Thus, the point of critique appears to be justified in that sense that cognitive predictors do not account for variance in noncognitive criteria. However, the basic idea of this dissertation was to carefully match the predictor and the criterion space to obtain thorough validity evidence. Thus, based on the assumption that cognitive predictors are more likely to predict task performance, whereas noncognitive predictors are more useful to predict citizenship and counterproductive behavior (McHenry, Hough, Toquam, Hanson, & Ashworth, 1990), there is no theoretical underpinning for the criticism saying that aptitude related tests need to be predictive of noncognitive performance. Instead, the need to use noncognitive predictors when predicting citizenship and counterproductive behavior is underscored by this point of critique.

The last criticism summarized under this assertion refers to the fact that noncognitive constructs explain incremental validity when predicting aptitude related outcomes. Based on the assumption that standardized tests were designed to predict G-GPA, they are criticized for not explaining all variance in the criterion (see Table 5.1). However, Sackett and colleagues (2008) argue that it is very unlikely to identify the one and only determinant of a criterion. This is in line with the findings presented in Chapter 4; although personality traits conceptualized in the Big Five and HEXACO taxonomy are well-established predictors of counterproductive academic behavior, trait mindfulness also explains a significant share of variance. This finding is promising as mindfulness interventions might be helpful to reduce the occurrence of counterproductive behavior, but it does not automatically narrow the utility of personality traits when it comes to the prediction of counterproductive academic behavior. The same holds true for standardized test, only because study skills (Credé & Kuncel, 2008) or achievement motivation and academic self-efficacy (Robbins, et al., 2004) explain incremental validity when predicting students' grades, the importance of standardized tests should not be put into question.

A third assertion encompasses the idea that “the appearance of validity is due to SES” (Sackett et al., 2008; p. 216). This criticism is grounded on the assumption that students originating from wealthier backgrounds have more resources available allowing them to perform better on standardized tests without being truly more capable (see Table 5.1). This is in line with the idea that SES erroneously inflates both students’ performance on standardized tests and their grades; implying that the significant relationship between test scores and performance is artificial in nature and disappears when the influence of SES is controlled for. Given that this is an issue often raised in the context of standardized testing, I considered the influence of SES when investigating the predictive validity of the GRE. In doing so, parental SES has been included in the hierarchical regression analysis and partial correlations between the three GRE parts and G-GPA controlling the influence of parental SES have been calculated. The findings reported in Chapter 2 imply that the predictive validity of the GRE-AW was still statistically significant when the influence of parental SES was held stable. Furthermore, the partial correlations were only marginally reduced compared to the zero-order correlations between GRE results and G-GPA alone. This pattern of results is in line with the findings of a meta-analysis (Sackett et al., 2009) showing that admission tests do not lose their predictive power when the influence of SES is controlled for. Thus, meta-analytical evidence (Sackett et al., 2009) and the findings presented in Chapter 2, do not support the assertion that the relationship between GRE and G-GPA is only an artifact of SES.

Taken together, the aforementioned findings provide empirical support for the multidimensional conceptualization of academic performance. As described in Chapter 1, the aim of this dissertation was to provide new insights to graduate students’ academic performance by broadening the criterion domain. Whereas, empirical research conducted in the work context shows that task, citizenship and counterproductive performance are distinct (Dalal, 2005; Hoffman, Blair, Meriac, & Woehr, 2007; Motowildo & Van Scotter, 1994; Sackett, 2002), this assumption had been implicitly adopted for the educational context so far (Meriac, 2012; Zettler, 2011). Especially at the master level where students are about to enter a constantly changing labor market, universities are encouraged to consider academic performance in broader terms. Results presented in this dissertation indeed show distinct patterns of relationships between the predictor and criterion variables. Whereas the GRE is predictive of task performance operationalized in terms of grades, the PPI obviously is a better predictor of citizenship and counterproductive behavior compared to U-GPA and GRE. Furthermore, the significant relationship between trait mindfulness and counterproductive academic behavior underscores the need to carefully match the predictor and the criterion domain.

Practical Implications

Whereas the aforementioned theoretical implications are important for scholars in the field of I/O psychology, the practical implications might be of special interest for stakeholders being responsible for admission decision in the academic context.

The first practical implication is related to the selection of master students. Due to changes going along with the Bologna declaration the European higher education system has considerably changed. Before the comprehensive reform students inscribed for an entire study program. In case of successful completion they received a diploma which they could use for application on the job market. Today, the system has been split-up into a bachelor and a master phase. After students have obtained their Bachelor degree they are free to decide whether they want to apply for a master program. In the Netherlands, students are permitted to inscribe for a master program in case they remain at the same university where they received their Bachelor degree. In many countries there is a large discrepancy between bachelor and master program capacities allowing not all Bachelor graduates to inscribe for a master program. Furthermore, not all students with a Bachelor degree are necessarily capable of obtaining a Master degree. Therefore, universities face the need to select students on the transition from bachelor to master phase. However, a number of countries (e.g., Germany, The Netherlands) have not been able to gain much experience with student selection at the graduate level. A common practice namely to select applicants based on their U-GPA is problematic due to missing comparability between different grading systems (e.g., Netherlands: 1 (lowest) -10 (highest) vs. Germany: 1 (highest) - 4 (lowest)) and grading severity at different universities. Thus, although student's U-GPA is easily accessible and its use is cost-efficient, it might result in flawed and unfair selection decisions making it vulnerable to legal claims. In contrast, the use of a well-established standardized test such as the GRE reduces the risk of legal claims as it is an objective criterion easily comparable between different applicants.

Another practical implication deriving from this dissertation concerns the possibilities to reduce the occurrence of counterproductive academic behavior. Based on the assumption that past behavior is a valid predictor of future behavior, it might be an option to simply ask students whether they have ever engaged in counterproductive behavior before. However, especially in a high-stake testing situation it is very likely that students answer in a social desirable way. Therefore, results presented in Chapter 3 are promising as they have shown that former supervisor ratings are valid predictors of future counterproductive academic behavior. Students with a high PPI score are less likely to engage in counterproductive academic behavior. Thus, selecting students based on PPI results is one reasonable option to reduce the likelihood that students engage in counterproductive behavior. The same holds true with respect to conscientiousness and honesty-humility. Students scoring high on conscientiousness and honesty-humility are less likely to show counterproductive behavior. So in case universities have the possibility to incorporate a traditional personality inventory to their selection procedure, they should aim to select those students scoring high on conscientiousness and honesty-humility.

Although mindfulness was also a valid predictor of counterproductive academic behavior as presented in Chapter 4, universities are not encouraged to screen out students based on their trait mindfulness level. Instead, mindfulness trainings might be an option to further increase students' mindfulness level after they have been admitted. In doing so not only the likelihood that they engage in counterproductive behavior will be reduced. Mindfulness interventions also have the potential to increase students' working memory capacity and test performance (Mrazek, Franklin, Philipps, Baird, & Schooler, 2013).

Taken together, this dissertation provides a basis for master students' assessment in the European context. The use of sound selection instruments is highly recommended as both parties lose in case students are either not capable or not willing to finish their master program. Whereas students often suffer from the experience of dropping-out of their study program, universities especially worry about wasting money and resources in case students leave before successfully graduating. Thus, by providing a demanding and tailored selection procedure universities can help applicants to get a realistic idea about the requirements they need to fulfill and encourage them to make a distinct decision with respect to their master program choice. At the same time, the use of noncognitive predictors allows universities to select applicants with relatively lower GRE scores in case this lack is outweighed by high noncognitive skills. In doing so, universities overcome the risk of admitting an overly homogenous group of students and foster creativity, innovation and problem solving instead. Especially on the graduate level where students are about to enter job market, they can benefit from experiences they make in diverse group settings. Finally, the idea to integrate mindfulness trainings in everyday students' life in order to potentially reduce the occurrence of counterproductive academic behavior goes along with several other advantages. Especially since the educational system has been split up into a Bachelor and a Master cycle students report suffering from a high stress level. Providing students with strategies to deal with these high demands in the academic context might help them from a professional, but also from a personal perspective. In the long run, universities offering mindfulness trainings might contribute to students' sustainable employment as they provide them with helpful coping mechanisms applicable in all kinds of stressful situation

Limitations and Future Research

Although providing important insights, the present dissertation is certainly not without limitations. In the remainder, I critically discuss these limitations and identify potential areas of future research.

One limitation exists with respect to the sample size. Especially with respect to Study 2 in which the predictive validity of the GRE has been investigated it would have been desirable to conduct a large-scale study. The initial idea of making participation obligatory for all new master students could unfortunately not be put into practice. Thus, international master students from five different faculties (School of Business and Economics, Faculty of Neuroscience, Faculty of Law, Faculty of Health, Medicine and Life Science, and Faculty of Arts and Social Science) were invited to participate in the study. Although the final sample size was comparably low ($N = 282$), the composition of the sample is almost equally distributed with respect to gender (55% female; 45% male) and representative of the diverse student body enrolled at Maastricht University (participants owned citizenship of 39 different countries) resembling its strong international orientation. However, it cannot be ruled out that some self-selection bias has influenced the composition of the sample. It might be that especially those students feeling capable with respect to standardized tests participated in our study or alternatively only those students feeling incompetent participated in order to gain some experience with cognitive loaded tests. Although this does not necessarily distort the predictive validity evidence found in Study 2, it might have led to restricted variance in our sample reducing the magnitude of the relationship between GRE and G-GPA. Thus, researchers aiming to replicate the findings of the present study should conduct large-scale studies and meta-analyses in order to increase the likelihood of making generalizable conclusions based on representative samples. With respect to Study 3, less than half of the students ($n = 115$) obtained PPI evaluations from former supervisors. The reasons for that are twofold. First, some students reported not to have a former supervisor knowing them well enough to provide an evaluation. Second, although those supervisors who did not accept the initial PPI request received two reminders some of them did not respond. Again the predictive validity evidence found in the present dissertation is probably not flawed due to the composition of the sample; however future research should try to find possibilities to obtain PPI evaluations also from those supervisors that are less committed to fill in the PPI in order to increase accuracy of the validity evidence.

A second limitation exists with respect to the measurement of university citizenship and counterproductive academic behavior in Study 3 (both) and Study 4 (only counterproductive academic behavior). For citizenship behavior, the German scale of Gehring (2006) was most suitable for our purposes. It has been used in a previous study showing sound psychometric properties (Zettler, 2011). We translated the scale into English by using a translation-back translation process and conducted a pilot study with 113 undergraduate students. In evaluating results of the pilot study several items were dropped based on both factor loadings and content related reasons. For instance the item "I am actively engaged in ensuring students' common interest (e.g., demonstrations against tuition fees.)" is not suitable for a population of Dutch students, as it refers to the German context where students demonstrated when tuition fees were introduced a couple of years ago. To measure counterproductive behavior, we agreed upon a well-established scale of workplace counterproductive behavior (Bennett & Robinson, 2000); we selected those items that could easily be adapted to the academic context (e.g., "Spent too much time fantasizing or dreaming instead of working." or "Made fun of someone at work.") and dropped those that were work specific (e.g., "Discussed confidential company information with an unauthorized person." or "Falsified a receipt to get reimbursed for more money than you spent on business expenses."). The 8-item scale was also pretested in the same pilot study mentioned above. Considering the psychometric properties of both scales presented in Study 3 and Study 4, citizenship and counterproductive behavior were both measured adequately. However, future research might consider replicating the findings of the present study by using different measures of citizenship and counterproductive behavior.

A third limitation concerns the operationalization of students' U-GPA. First, students were asked to indicate their U-GPA. Thus, the accuracy of the information is not verified. Furthermore, the conversion of students' U-GPA into the Dutch grading system was not without ambiguity. As mentioned before, the sample consisted of students who obtained their Bachelor degrees in 35 different countries. Based on different conversion schemes, a research assistant and I converted the grades into the Dutch grading system. In case the discrepancy between the two conversions was not larger than one grade, a mean score was calculated. In case the discrepancy was larger than one grade, the converted grades were reexamined more in detail. However, given that conversion of grades into the Dutch grading system was not without difficulties, I ran the main analyses investigating the incremental validity of the GRE and the PPI above and beyond U-GPA by using a subsample of students who all received their Bachelor degree at Dutch universities (see Appendix 1 - 3). Whereas for the PPI the same pattern of results emerged in the subsample as in the overall sample, the incremental validity of the GRE decreased. Thus, both research and practice might profit from better standards to convert grades from one grading system to another.

A final limitation exists regarding the generalizability of the results. First, Maastricht University is a fully-bilingual (Dutch & English) university attracting a diverse student body. As mentioned in Chapter 1 of this dissertation, the validity of the GRE for non-native English speakers has been described as a concern which might limit generalizability of validity evidence that has been gained in the United States (Kuncel et al., 2001). It might be that students' English proficiency moderates the relationship between GRE scores and subsequent study success. In general, Dutch people are considered to have a very high English proficiency level (EF - English Proficiency Index, 2014). The situation differs in other European countries (e.g., France, Italy or Spain) in which people are considered to have a moderate English proficiency level (EF - English Proficiency Index, 2014) and in which only a few universities offer English speaking master programs. In these countries, universities attract a higher percentage of domestic students that might have a lower English proficiency level compared to the students enrolled at Maastricht University. Thus, in order to investigate whether the validity evidence provided in Study 2 is generalizable throughout Europe, future research needs to be conducted at other European universities where students' English proficiency might be more of a concern. Second, Maastricht University is characterized by a problem based learning (PBL) teaching system in which students learn in small, self-regulated student groups of 10-14 persons. The PBL system is clearly student centered. Students work together in small groups in order to solve practical or academic problems. Thus, proper group functioning in the PBL system is dependent on team processes, such as communication, knowledge sharing, and collaboration between students. In consequence, noncognitive performance aspects like citizenship and counterproductive behavior are critically important for the functioning of these groups. Thus, whereas Maastricht University is especially suitable to investigate the usefulness of the PPI (Chapter 3) and trait mindfulness (Chapter 4) to predict citizenship and counterproductive behavior, validity evidence might differ at other European Universities where the natural occurrence of citizenship and counterproductive behavior is reduced. Thus, future research focusing on other educational systems (i.e., mainly offering lectures or web based distance education) needs to be conducted in order to investigate whether predictive value of the PPI and mindfulness can be generalized across different educational systems in Europe.

Conclusion

This dissertation adds to the literature on cognitive and noncognitive predictors of study success by demonstrating that the GRE predicts study success operationalized in terms of grades whereas the PPI predicts university citizenship and counterproductive academic behavior. Assuming that the consequences of counterproductive academic behavior are especially severe from students' and universities' perspective, this dissertation sets out to establish the link between trait mindfulness and counterproductive academic behavior. Thus, increasing trait mindfulness through the use of training interventions after students have been already admitted might be a promising mechanism through which the occurrence of counterproductive behavior could be reduced. Taken together, this dissertation underscores the need to conceptualize academic performance in broader terms and to match the predictor and the criterion space when investigating the validity of cognitive and noncognitive assessment and development tools. It thereby contributes to the development of a student body having the required knowledge and skills to be successful in the academic context while being well-prepared for their future careers.

References

- Aamodt, M. G. (1999). *Applied industrial organizational psychology* (3rd ed.). London: Brooks/Cole.
- Aamodt, M. G., Bryan, D. A., & Whitcomb, A. J. (1993). Predicting performance with letters of recommendation. *Public Personnel Management*, 22, 81–90. doi: 10.1177/009102609302200106
- Aguinis, H. (1995). Statistical power with moderated multiple regression in management research. *Journal of Management*, 21, 1141–1158. doi: 10.1177/014920639502100607
- Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks, CA: Sage.
- Allison, B. J., Voss, R. S., & Dryer, S. (2001). Student classroom and career success: The role of organizational citizenship behavior. *Journal of Education for Business*, 76, 282–288. doi: 10.1080/08832320109599650
- Ashton, M. C., & Lee, K. (2005). Honesty–Humility, the Big Five, and the Five–Factor Model. *Journal of Personality*, 73, 1321–1354. doi: 10.1111/j.1467-6494.2005.00351.x
- Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, 11, 150–166. doi: 10.1177/1088868306294907
- Ashton, M. C., & Lee, K. (2009). The HEXACO–60: A short measure of the major dimensions of personality. *Journal of Personality Assessment*, 91, 340–345. doi: 10.1080/00223890902935878
- Baer, R. A., Fischer, S., & Huss, D. B. (2005). Mindfulness and acceptance in the treatment of disordered eating. *Journal of Rational-Emotive and Cognitive-Behavior Therapy*, 23, 281–300. doi: 10.1007/s10942-005-0015-9
- Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, 13, 27–45. doi: 10.1177/1073191105283504
- Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta–analysis. *Personnel Psychology*, 44, 1–26. doi: 10.1111/j.1744-6570.1991.tb00688.x
- Barron, J. (2013, March 6). On Campus, Costly Target of Brazen Thefts: Nutella. *The New York Times*. Retrieved from http://www.nytimes.com/2013/03/07/nyregion/for-columbia-students-nutella-in-a-dining-hall-may-be-too-tempting.html?partner=rss&emc=rss&_r=3&
- Becker, R., & Kolster, R. (2012, January). *International student recruitment: Policies and developments in selected countries*. The Hague, The Netherlands: NUFFIC (Netherlands Organisation for International Cooperation in Higher Education). Retrieved from <https://www.nuffic.nl/en/library/international-student-recruitment.pdf>
- Bennett, R. J., & Robinson, S. L. (2000). Development of a measure of workplace deviance. *Journal of Applied Psychology*, 85, 349–360. doi: 10.1037/0021-9010.85.3.349
- Bentler, P. M. (2009). Alpha, dimension-free, and model-based internal consistency reliability. *Psychometrika*, 74, 137–143. doi: 10.1007/s11336-008-9100-1
- Berry, C. M., Carpenter, N. C., & Barratt, C. L. (2012). Do other-reports of counterproductive work behavior provide an incremental contribution over self-reports? A meta-analytic comparison. *Journal of Applied Psychology*, 97, 613–636. doi: 10.1037/a0026739

References

- Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal deviance, organizational deviance, and their common correlates: A review and meta-analysis. *Journal of Applied Psychology, 92*, 410–424. doi: 10.1037/0021-9010.92.2.410
- Blankenship, K. L., & Whitley, B. E. (2000). Relation of general deviance to academic dishonesty. *Ethics & Behavior, 10*, 1–12. doi: 10.1207/S15327019EB1001_1
- Bleske-Rechek, A., & Browne, K. (2014). Trends in GRE scores and graduate enrollments by gender and ethnicity. *Intelligence, 46*, 25–34. doi: 10.1016/j.intell.2014.05.005
- Bohlmeijer, E., Prenger, R., Taal, E., & Cuijpers, P. (2010). The effects of mindfulness-based stress reduction therapy on mental health of adults with a chronic medical disease: A meta-analysis. *Journal of Psychosomatic Research, 68*, 539–544. doi: 10.1016/j.jpsychores.2009.10.005
- Borman, W. C., Brantley, L. B., & Hanson, M. A. (2014). Progress toward understanding the structure and determinants of job performance: A focus on task and citizenship performance. *International Journal of Selection and Assessment, 22*, 422–431. doi: 10.1111/ijsa.12088
- Borman, W. C., & Motowildo, S. M. (1993). Expanding the criterion domain to include elements of contextual performance. Chapter in N. Schmitt and W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 71–98). San Francisco: Jossey-Bass
- Borman, W. C., & Motowildo, S. J. (1997). Task performance and contextual performance: The meaning for personnel selection research. *Human Performance, 10*, 99–109. doi: 10.1207/s15327043hup1002_3
- Bridgeman, B., Burton, N., & Cline, F. (2009). A note on presenting what predictive validity numbers mean. *Applied Measurement in Education, 22*, 109–119. doi: 10.1080/08957340902754577
- Briel, J., Bejar, I., Chandler, M., Powell, G., Manning, K., Robinson, D., Smallwood, T., Vitella, S., & Welsh, C. (2000). GRE Horizons Planning Initiative. (Unpublished Graduate Record Examination report). Princeton, NJ: Educational Testing Service.
- Brown, B. (2007). Letters: The utility of standardized tests. *Science, 316*, 1694–1695. Retrieved from <http://www.evsc.virginia.edu/wp-content/uploads/2007-LerdauAvery-Science.pdf>
- Brown, K. W., & Ryan, R. M. (2003). The benefits of being present: mindfulness and its role in psychological well-being. *Journal of Personality and Social Psychology, 84*, 822–848. doi: 10.1037/0022-3514.84.4.822
- Buchmann, C., & Dalton, B. (2002). Interpersonal influences and educational aspirations in 12 countries: The importance of institutional context. *Sociology of Education, 75*, 99–122. doi: 10.2307/3090287
- Caldwell, K., Harrison, M., Adams, M., Quin, R., & Greeson, J., (2010). Developing mindfulness in college students through movement-based courses: Effects on self-regulatory self-efficacy, mood, stress, and sleep quality. *Journal of American College Health, 58*, 433–442. doi: 10.1080/07448480903540481
- Carmody, J., & Baer, R. A. (2008). Relationships between mindfulness practice and levels of mindfulness, medical and psychological symptoms and well-being in a mindfulness-based stress reduction program. *Journal of Behavioral Medicine, 31*, 23–33. doi: 10.1007/s10865-007-9130-7

- Carpenter, N. C., Berry, C. M., & Houston, L. (2014). A meta-analytic comparison of self-reported and other-reported organizational citizenship behavior. *Journal of Organizational Behavior*, 35, 547–574. doi: 10.1002/job.1909
- Chapman, B. P., & Hayslip, Jr, B. (2005). Incremental validity of a measure of emotional intelligence. *Journal of Personality Assessment*, 85, 154–169. doi: 10.1207/s15327752jpa8502_08
- Collard, P., Avny, N., & Boniwell, I., (2008). Teaching mindfulness based cognitive therapy (MBCT) to students: The effects of MBCT on the levels of mindfulness and subjective well-being. *Counselling Psychology Quarterly*, 21, 323–336. doi: 10.1080/09515070802602112
- Connelly, B. S., & Hülshager, U. R. (2012). A narrower scope or a clearer lens for personality? Examining sources of observers' advantages over self-reports for predicting performance. *Journal of Personality*, 80, 603–631. doi: 10.1111/j.1467-6494.2011.00744.x
- Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin*, 136, 1092–1122. doi: 10.1037/a0021212
- Costa, P. T. Jr., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual. Odessa, FL: Psychological Assessment Resources.
- Côté, S., & Miners, C. T. (2006). Emotional intelligence, cognitive intelligence, and job performance. *Administrative Science Quarterly*, 51, 1–28. Retrieved from <http://www.jstor.org/stable/20109857>
- Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science*, 3, 425–453. doi: 10.1111/j.1745-6924.2008.00089.x
- Credé, M., & Niehorster, S. (2009). Individual difference influences on self-focused and other-focused counterproductive student behaviors. *Personality and Individual Differences*, 47, 769–776. doi: 10.1016/j.paid.2009.06.018
- Creswell, J. D., Way, B. M., Eisenberger, N. I., & Lieberman, M. D. (2007). Neural correlates of dispositional mindfulness during affect labeling. *Psychosomatic Medicine*, 69, 560–565. doi: 10.1097/PSY.0b013e3180f6171f
- Crosby, F. J., Iyer, A., Clayton, S., & Downing, R. A. (2003). Affirmative action: Psychological data and the policy debates. *American Psychologist*, 58, 93–115. doi: 10.1037/0003-066X.58.2.93
- Dalal, R. S. (2005). A meta-analysis of the relationship between organizational citizenship behavior and counterproductive work behavior. *Journal of Applied Psychology*, 90, 1241–1255. doi: 10.1037/0021-9010.90.6.1241
- Dane, E. (2011). Paying attention to mindfulness and its effects on task performance in the workplace. *Journal of Management*, 37, 997–1018. doi: 10.1177/0149206310367948
- de Vries, A., de Vries, R. E., & Born, M. P. (2011). Broad versus narrow traits: Conscientiousness and Honesty–Humility as predictors of academic criteria. *European Journal of Personality*, 25, 336–348. doi: 10.1002/per.795
- Dörnfelder, A. (2010, August 31). Wenn der Abschluss an der Top-Uni zum Nachteil wird. [If graduating at a top-university becomes a disadvantage]. *Handelsblatt*. Retrieved from: <http://www.handelsblatt.com/politik/oekonomie/nachrichten/das-master-desaster-wenn-der-abschluss-an-der-top-uni-zum-nachteil-wird-seite-all/3527486-all.html>

References

- Education, Audiovisual and Culture Executive Agency (2010). Focus on higher education in Europe 2010: The impact of the Bologna process. Brussels, Belgium: Education, Audiovisual and Culture Executive Agency. Retrieved from http://eacea.ec.europa.eu/education/eurydice/documents/thematic_reports/122EN.pdf
- Enright, M. K., & Gitomer, D. (1989). Toward a Description of Successful Graduate Students. (GREB 85-17R; RR 89-09). Princeton, NJ: Educational Testing Service. Retrieved from <http://origin-www.ets.org/Media/Research/pdf/RR-89-09-Enright.pdf>
- Farsides, T., & Woodfield, R. (2003). Individual differences and undergraduate academic success: The roles of personality, intelligence, and application. *Personality and Individual Differences*, 34, 1225-1243. doi: 10.1016/S0191-8869(02)00111-3
- Fox, S., Spector, P. E., Goh, A., & Bruursema, K. (2007). Does your coworker know what you're doing? Convergence of self-and peer-reports of counterproductive work behavior. *International Journal of Stress Management*, 14, 41-60. doi: 10.1037/1072-5245.14.1.41
- Gehring, F. (2006). University Citizenship Behavior. Development and validation of an inventory on contextual performance at universities (Unpublished diploma thesis). RWTH Aachen University, Aachen, Germany.
- Giluk, T. L. (2009). Mindfulness, Big Five personality, and affect: A meta-analysis. *Personality and Individual Differences*, 47, 805-811. doi: 10.1016/j.paid.2009.06.026
- Glomb, T. M., Duffy, M. K., Bono, J. E., & Yang, T. (2011). Mindfulness at work. In J. Martocchio, H. Liao, & A. Joshi (Eds.), *Research in Personnel and Human Resource Management* (pp. 115-157).
- Goldberg, E. L., & Alliger, G. M. (1992). Assessing the validity of the GRE for students in psychology: A validity generalization approach. *Educational and Psychological Measurement*, 52, 1019-1027. doi: 10.1177/0013164492052004026
- Gonzalez-Mulé, E., Mount, M. K., & Oh, I. (2014). A meta-analysis of the relationship between general mental ability and nontask performance. *Journal of Applied Psychology*, 99, 1222-1243. doi: 10.1037/a0037547
- Gore, P. A. (2006). Academic self-efficacy as a predictor of college outcomes: Two incremental validity studies. *Journal of Career Assessment*, 14, 92-115. doi: 10.1177/1069072705281367
- Gore, J. S., Kiefner, A. E., & Combs, K. M. (2012). Personality traits that predict academic citizenship behavior. *Journal of Applied Social Psychology*, 42, 2433-2456. doi: 10.1111/j.1559-1816.2012.00948.x
- Gruys, M. L., & Sackett, P. R. (2003). Investigating the dimensionality of counterproductive work behavior. *International Journal of Selection and Assessment*, 11, 30-42. doi: 10.1111/1468-2389.00224
- Hakstian, A. R., Farrell, S., & Tweed, R. G. (2002). The assessment of counterproductive tendencies by means of the California Psychological Inventory. *International Journal of Selection and Assessment*, 10, 58-86. doi: 10.1111/1468-2389.00194
- Hamilton, L. C. (1992). *Regression with graphics: A second course in applied statistics* (Vol. 1, No. 1). Belmont, CA: Duxbury press.
- Harding, T. S., Carpenter, D. D., Finelli, C. J., & Passow, H. J. (2004). Does academic dishonesty relate to unethical behavior in professional practice? An exploratory study. *Science and Engineering Ethics*, 10, 311-324. doi: 10.1007/s11948-004-0027-3

- Heinz, W. R. (1999). *From education to work: Cross national perspectives*. Cambridge University Press.
- Henz, U. (1997). Der nachgeholte Erwerb allgemeinbildender Schulabschlüsse. Analysen zur quantitativen Entwicklung und sozialen Selektivität. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 49, 223–241.
- Hell, B., Trapmann, S. & Schuler, H. (2007). Eine Meta-Analyse der Validität von fachspezifischen Studierfähigkeitstests. [A meta-analytic investigation of subject-specific admission tests in German-speaking countries.] *Empirische Pädagogik*, 21, 251–270. Retrieved from <http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-71901>
- Hoffman, B. J., Blair, C. A., Meriac, J. P., & Woehr, D. J. (2007). Expanding the criterion domain? A quantitative review of the OCB literature. *Journal of Applied Psychology*, 92, 555–566. doi: 10.1037/0021-9010.92.2.555
- Hofmann, S. G., Sawyer, A. T., Witt, A. A., & Oh, D. (2010). The effect of mindfulness-based therapy on anxiety and depression: A meta-analytic review. *Journal of Consulting and Clinical Psychology*, 78, 169–183. doi: 10.1037/a0018555
- Hough, L. M., Oswald F., & Ployhart R. (2001). Determinants, detection, and amelioration of adverse impact in personnel selection procedures: Issues, evidence, and lessons learned. *International Journal of Selection and Assessment*, 9, 152–194. doi: 10.1111/1468-2389.00171
- Howell, A. J., & Buro, K. (2011). Relations among mindfulness, achievement-related self-regulation, and achievement emotions. *Journal of Happiness Studies*, 12, 1007–1022. doi: 10.1007/s10902-010-9241-7
- Hülshager, U. R., Alberts, H. J., Feinholdt, A., & Lang, J. W. (2013). Benefits of mindfulness at work: The role of mindfulness in emotion regulation, emotional exhaustion, and job satisfaction. *Journal of Applied Psychology*, 98, 310–325. doi: 10.1037/a0031313
- Hülshager, U. R., Maier, G. W., & Stumpp, T. (2007). Validity of general mental ability for the prediction of job performance and training success in Germany: A meta-analysis. *International Journal of Selection and Assessment*, 15, 3–18. doi: 10.1111/j.1468-2389.2007.00363.x
- Hunsley, J., & Meyer, G. J. (2003). The incremental validity of psychological testing and assessment: Conceptual, methodological, and statistical issues. *Psychological Assessment*, 15, 446–455. doi: 10.1037/1040-3590.15.4.446
- Hunter, J. E., & Schmidt, F. L. (Eds.). (2004). *Methods of meta-analysis: Correcting error and bias in research findings* (2nd ed.). Thousand Oaks, CA: Sage.
- Johnson, V. E. (2003). *Grade inflation: A crisis in college education*. New York: Springer.
- Kabat-Zinn, J. (1982). An outpatient program in behavioral medicine for chronic pain patients based on the practice of mindfulness meditation: Theoretical considerations and preliminary results. *General Hospital Psychiatry*, 4, 33–47. doi: 10.1016/0163-8343(82)90026-3
- Kabat-Zinn, J. (2003). Mindfulness-based interventions in context: past, present, and future. *Clinical Psychology: Science and Practice*, 10, 144–156. doi: 10.1093/clipsy.bpg016
- Kerckhoff, A. C. (2001). Education and social stratification processes in comparative perspective. *Sociology of Education*, 3–18. doi: 10.2307/2673250
- Kim, S., & Kyllonen, P. C. (2007). Rasch measurement in developing faculty ratings of students applying to graduate school. *Journal of Applied Measurement*, 9, 168–181. Retrieved from <http://europepmc.org/abstract/MED/18480513>
- Kohn, A. (2001). Two cheers for an end to the SAT. *Chronicle of Higher Education*, 9, p. B–12.

References

- Koppenhaver, G. D. (2006). Absent and accounted for: Absenteeism and cooperative learning. *Decision Sciences Journal of Innovative Education*, 4, 29–49. doi: 10.1111/j.1540-4609.2006.00100.x
- Krishnakumar, S., & Robinson, M. D. (2015). Maintaining an even keel: An affect-mediated model of mindfulness and hostile work behavior. *Emotion*, 1–11. doi: 10.1037/emo0000060
- Kuncel, N. R., Credé, M., & Thomas, L. L. (2007). A comprehensive meta-analysis of the predictive validity of the Graduate Management Admission Test (GMAT) and undergraduate grade point average (UGPA). *Academy of Management Learning and Education*, 6, 51–68. doi: 10.5465/AMLE.2007.24401702
- Kuncel, N.R., & Hezlett, S.A. (2007). Standardized tests predict graduate students' success. *Science*, 315, 1080–1081. doi: 10.1126/science.1136618
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, 127, 162–181. doi: 10.1037/0033-2909.127.1.162
- Kuncel, N. R., & Hezlett, S. A. (2010). Fact and fiction in cognitive ability testing for admissions and hiring decisions. *Current Directions in Psychological Science*, 19, 339–345. doi: 10.1177/0963721410389459
- Kuncel, N. R., Wee, S., Serafin, L., & Hezlett, S. A. (2010). The validity of the Graduate Record Examination for master's and doctoral programs: A meta-analytic investigation. *Educational and Psychological Measurement*, 70, 340–352. doi: 10.1177/0013164409344508
- Kyllonen, P. C. (2008). The research behind the ETS Personal Potential Index (PPI). Princeton, NY, ETS. Retrieved from http://www.ets.org/Media/Products/PPI/10411_PPI_bkgrd_report_RD4.pdf
- Kyllonen, P. C., & Kim, S. (2005, April). Personal qualities in higher education: Dimensionality of faculty ratings of students applying to graduate school. Paper presented at the annual meeting of the American Educational Research Association, Montréal, Canada.
- Kyllonen, P., Walters, A. M., & Kaufman, J. C. (2005). Noncognitive constructs and their assessment in graduate education: A review. *Educational Assessment*, 10, 153–184. doi: 10.1207/s15326977ea1003_2
- Lahey, C. E., Campbell, W. K., Brown, K. W., & Goodie, A. S. (2007). Dispositional mindfulness as a predictor of the severity of gambling outcomes. *Personality and Individual Differences*, 43, 1698–1710. doi: 10.1016/j.paid.2007.05.007
- Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research*, 39, 329–358. doi: 10.1207/s15327906mbr3902_8
- Lee, K., Ashton, M.C., & de Vries, R.E. (2005a). Predicting workplace delinquency and integrity with the HEXACO and Five-Factor models of personality structure. *Human Performance*, 18, 179–197. doi: 10.1207/s15327043hup1802_4
- Lee, K., Ashton, M.C., & Shin, K.H. (2005b). Personality correlates of workplace anti-social behavior. *Applied Psychology: An International Review*, 54, 81–98. doi: 10.1111/j.1464-0597.2005.00197.x

- LePine, J. A., Erez, A., & Johnson, D. E. (2002). The nature and dimensionality of organizational citizenship behavior: A critical review and meta-analysis. *Journal of Applied Psychology*, 87, 52–65. doi: 10.1037/0021-9010.87.1.52
- Lievens, F., Buyse, T., & Sackett, P. R. (2005). The operational validity of a video-based situational judgment test for medical college admissions: Illustrating the importance of matching predictor and criterion construct domains. *Journal of Applied Psychology*, 90, 442–452. doi: 10.1037/0021-9010.90.3.442
- Lievens, F., & Coetsier, P. (2002). Situational tests in student selection: An examination of predictive validity, adverse impact, and construct validity. *International Journal of Selection and Assessment*, 10, 245–257. doi: 10.1111/1468-2389.00215
- Lievens, F., & Sackett, P. R. (2006). Video-based versus written situational judgment tests: A comparison in terms of predictive validity. *Journal of Applied Psychology*, 91, 1181–1188. doi: 10.1037/0021-9010.91.5.1181
- Lievens, F., & Sackett, P. R. (2012). The validity of interpersonal skills assessment via situational judgment tests for predicting academic success and job performance. *Journal of Applied Psychology*, 97, 460–468. doi: 10.1037/a0025741
- Liu, O. L., Minsky, J., Ling, G., & Kyllonen, P. (2009). Using the standardized letters of recommendation in selection: Results from a multidimensional Rasch model. *Educational and Psychological Measurement*, 69, 475–492. doi: 10.1177/0013164408322031
- Looker, E. D. (1989). Accuracy of proxy reports of parental status characteristics. *Sociology of Education*, 257–276. Retrieved from <http://www.jstor.org/stable/2112830>
- Lutz, J., Herwig, U., Opialla, S., Hittmeyer, A., Janke, L., Rufer, M., et al. (2014). Mindfulness and emotion regulation—An fMRI study. *Social Cognitive and Affective Neuroscience*, 9, 776–785. doi: 10.1093/scan/nst043
- Marcus, B., Goffin, R. D., Johnston, N. G., & Rothstein, M. G. (2007). Personality and cognitive ability as predictors of typical and maximum managerial performance. *Human Performance*, 20, 275–285. doi: 10.1080/08959280701333362
- Marcus, B., Lee, K., & Ashton, M. C. (2007). Personality dimensions explaining relationships between integrity tests and counterproductive behavior: Big Five, or one in addition?. *Personnel Psychology*, 60, 1–34. doi: 10.1111/j.1744-6570.2007.00063.x
- Marcus, B., & Schuler, H. (2004). Antecedents of counterproductive behavior at work: A general perspective. *Journal of Applied Psychology*, 89, 647–660. doi: 10.1037/0021-9010.89.4.647
- Marcus, B., Taylor, O. A., Hastings, S. E., Sturm, A., & Weigelt, O. (2013). The structure of counterproductive work behavior a review, a structural meta-analysis, and a primary study. *Journal of Management*, 0149206313503019. doi: 10.1177/0149206313503019
- Marquardt, D. W. (1980). You should standardize the predictor variables in your regression models. *Journal of the American Statistical Association*, 75, 87–91. doi: 10.1080/01621459.1980.10477430
- Matta, F. K., Erol-Korkmaz, H. T., Johnson, R. E., & B-çaks-z, P. (2014). Significant work events and counterproductive work behavior: The role of fairness, emotions, and emotion regulation. *Journal of Organizational Behavior*, 35, 920–944. doi

References

- Mazzarol, T., & Soutar, G. N. (2002). "Push-pull" factors influencing international student destination choice. *International Journal of Educational Management*, 16, 82-90. doi: 10.1108/09513540210418403
- McCabe, D. L., Butterfield, K. D., & Trevino, L. K. (2006). Academic dishonesty in graduate business programs: Prevalence, causes, and proposed action. *Academy of Management Learning and Education*, 5, 294-305. doi: 10.5465/AMLE.2006.22697018
- McCarthy, J. M., & Goffin, R. D. (2001). Improving the validity of letters of recommendation: An investigation of three standardized reference forms. *Military Psychology*, 13, 199-222. doi: 10.1207/S15327876MP1304_2
- McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L., III (2007). Situational judgment tests, response instructions and validity: A meta-analysis. *Personnel Psychology*, 60, 63-91. doi: 10.1111/j.1744-6570.2007.00065.x
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology*, 86, 730-740. doi: 10.1037/0021-9010.86.4.730
- McDaniel, M. A., & Nguyen N.T. (2001). Situational judgment tests: A review of practice and constructs assessed. *International Journal of Selection and Assessment*, 9, 103-113. doi: 10.1111/1468-2389.00167.
- McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashforth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology*, 43, 335-354. doi: 10.1111/j.1744-6570.1990.tb01562.x
- McManus, M. A., & Kelly, M. L. (1999). Personality measures and biodata: Evidence regarding their incremental predictive value in the life insurance industry. *Personnel Psychology*, 52, 137-148. doi: 10.1111/j.1744-6570.1999.tb01817.x
- Meriac, J. P. (2012). Work ethic and academic performance: Predicting citizenship and counterproductive behavior. *Learning and Individual Differences*, 22, 549-553. doi: 10.1016/j.lindif.2012.03.015
- Morphew, C. C., & Hartley, M. (2006). Mission statements: A thematic analysis of rhetoric across institutional type. *The Journal of Higher Education*, 77, 456-471. doi: 10.1353/jhe.2006.0025
- Morrison, T., & Morrison, M. (1995). A meta-analytic assessment of the predictive validity of the quantitative and verbal components of the Graduate Record Examination with graduate grade point average representing the criterion of graduate success. *Educational and Psychological Measurement*, 55, 309-316. doi: 10.1177/0013164495055002015
- Motowildo, S. J., Borman, W. C., & Schmit, M. J. (1997). A theory of individual differences in task and contextual performance. *Human Performance*, 10, 71-83. doi: 10.1207/s15327043hup1002_1
- Motowildo, S. J., & Van Scotter, J. R. (1994). Evidence that task performance should be distinguished from contextual performance. *Journal of Applied Psychology*, 79, 475-480. doi: 10.1037/0021-9010.79.4.475

- Mount, M. K., Witt, L. A., & Barrick, M. R. (2000). Incremental validity of empirically keyed biodata scales over GMA and the five factor personality constructs. *Personnel Psychology*, 53, 299–323. doi: 10.1111/j.1744-6570.2000.tb00203.x
- Mrazek, M. D., Franklin, M. S., Phillips, D. T., Baird, B., & Schooler, J. W. (2013). Mindfulness training improves working memory capacity and GRE performance while reducing mind wandering. *Psychological Science*, 24, 776–781. doi: 10.1177/0956797612459659
- Müller, W., & Shavit, Y. (1998). The institutional embeddedness of the stratification process. In Y. Shavit & W. Müller (Eds.), *From school to work* (pp. 1–48). Oxford: Clarendon Press.
- Ng, T. W., & Feldman, D. C. (2009). How broadly does education contribute to job performance? *Personnel Psychology*, 62, 89–134. doi: 10.1111/j.1744-6570.2008.01130.x
- Nguyen, N. T., Biderman, M. D., & McDaniel, M. A. (2005). Effects of response instructions on faking a situational judgment test. *International Journal of Selection and Assessment*, 13, 250–260. doi: 10.1111/j.1468-2389.2005.00322.x
- Nicklin, J. M., & Roch, S. G. (2009). Letters of recommendation: Controversy and consensus from expert perspectives. *International Journal of Selection and Assessment*, 17, 76–91. doi: 10.1111/j.1468-2389.2009.00453.x
- Nuffic (2006). *Cijfers ontcijferd: Notitie aangaande het omzetten van (examen) cijfers: een onderzoek naar de verdeling van cijfers en grades in het onderwijs van Nederland, de Verenigde Staten, Canada en het Verenigd Koninkrijk*. [Grade deciphered: A note regarding the transformation of exam grades: A study on the distribution of grades in higher education in the Netherlands, the United States, Canada, and the United Kingdom]. The Hague, NL: Author. Retrieved from <http://www.nuffic.nl/root/nuffic/bestanden/documenten/over-de-nuffic/publicaties/factsheet-cijfers-ontcijferd.pdf>
- Nuffic (2012). Country module – the Netherlands. Degrees and qualifications in the Netherlands, 3, 1–39. Retrieved from <https://www.nuffic.nl/en/library/country-module-netherlands.pdf>
- Oh, I. S., Lee, K., Ashton, M. C., & de Vries, R. E. (2011). Are dishonest extraverts more harmful than dishonest introverts? The interaction effects of honesty, humility and extraversion in predicting workplace deviance. *Applied Psychology: An International Review*, 60, 496–516. doi: 10.1111/j.1464-0597.2011.00445.x
- Oldfield, K., & Hutchinson, J. R. (1997). Predictive validity of the Graduate Record Examination with and without range restriction. *Psychological Reports*, 81, 211–220. doi: 10.2466/pr0.1997.81.1.211
- Organ, D. W. (1990). The motivational basis of organizational citizenship behavior. *Research in Organizational Behavior*, 12, 43–72.
- Organ, D. W. (1997). Organizational citizenship behavior: It's construct clean-up time. *Human Performance*, 10, 85–97. doi: 10.1207/s15327043hup1002_2

References

- Oswald, F. L., Schmitt, N., Kim, B. H., Ramsay, L. J., & Gillespie, M. A. (2004). Developing a biodata measure and situational judgment inventory as predictors of college student performance. *Journal of Applied Psychology*, 89, 187–207. doi: 10.1037/0021-9010.89.2.187
- Pérez-Peña, R. (2013, February 1). Students disciplined in Harvard scandal. *The New York Times*. Retrieved from http://www.nytimes.com/2013/02/02/education/harvard-forced-dozens-to-leave-in-cheating-scandal.html?_r=0.
- Ployhart, R. E., & Holtz, B. C. (2008). The diversity-validity dilemma: Strategies for reducing racioethnic and sex subgroup differences and adverse impact in selection. *Personnel Psychology*, 61, 153–172. doi: 10.1111/j.1744-6570.2008.00109.x
- Poropat, A. E. (2009). A meta-analysis of the Five-Factor Model of personality and academic performance. *Psychological Bulletin*, 135, 322–338. doi: 10.1037/a0014996
- Poropat, A. E. (2011). The role of citizenship performance in academic achievement and graduate employability. *Education and Training*, 53, 499–514. doi: 10.1108/00400911111159467
- Powers, D. E. (2004). Validity of Graduate Record Examinations (GRE) General Test Scores for Admissions to Colleges of Veterinary Medicine. *Journal of Applied Psychology*, 89, 208–219. doi: 10.1037/0021-9010.89.2.208
- Ramsay, L., Schmitt, N., Oswald, F. L., Kim, B. H., & Gillespie, M. (2006). The impact of situational context variables on responses to biodata and situational judgment inventory items. *Psychology Science*, 48, 268–287. Retrieved from <http://io.psy.msu.edu/cbstudy/Ramsay%20SJI%20-%20Psychology%20Sci%202006.pdf>
- Ramsburg, J. T., & Youmans, R. J. (2013). Meditation in the higher-education classroom: Meditation training improves student knowledge retention during lectures. *Mindfulness*, 1–11. doi: 10.1007/s12671-013-0199-5
- Range, L. M., Menyhert, A., Walsh, M. L., Hardin, K. N., Ellis, J. B. & Craddick, R. (1991) Letters of recommendation: Perspectives, recommendations, and ethics. *Professional Psychology: Research and Practice*, 22, 389–392. doi: 10.1037/0735-7028.22.5.389
- Regehr, C., Glancy, D., & Pitts, A. (2013). Interventions to reduce stress in university students: A review and meta-analysis. *Journal of Affective Disorders*, 148, 1–11. doi: 10.1016/j.jad.2012.11.026
- Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijsma. *Psychometrika*, 74, 145–154. doi: 10.1007/s11336-008-9102-z
- Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, 138, 353–387. doi: 10.1037/a0026838
- Ridgell, S., & Lounsbury, J. W. (2004). Predicting collegiate academic success: General intelligence, "Big Five" personality traits, and work drive. *College Student Journal*, 38, 607–618. Retrieved from <http://www.jcsdonline.org/>
- Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, 130, 261–288. doi: 10.1037/0033-2909.130.2.261

- Roberts, B. W., Harms, P. D., Caspi, A., & Moffitt, T. E. (2007). Predicting the counterproductive employee in a child-to-adult prospective study. *Journal of Applied Psychology, 92*, 1427–1436. doi: 10.1037/0021-9010.92.5.1427
- Robinson, S. L., & Bennett, R. J. (1995). A typology of deviant workplace behaviors: A multidimensional scaling study. *Academy of Management Journal, 38*, 555–572. doi: 10.2307/256693
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*, 1–36. Retrieved from http://www.lce.esalq.usp.br/arquivos/aulas/2013/encontro_ppg/Lucia/paper.pdf
- Rotundo, M., & Sackett, P. R. (2002). The relative importance of task, citizenship, and counterproductive performance for supervisor ratings of overall performance: A policy capturing study. *Journal of Applied Psychology, 87*, 66–80. doi: 10.1177/014920639402000307
- Sackett, P. R. (2002). The structure of counterproductive work behaviors: Dimensionality and relationships with facets of job performance. *International Journal of Selection and Assessment, 10*, 5–11. doi: 10.1111/1468-2389.00189
- Sackett, P. R., Borneman, M. J., & Connelly, B. S. (2008). High stakes testing in higher education and employment: Appraising the evidence for validity and fairness. *American Psychologist, 63*, 215–227. doi: 10.1037/0003-066X.56.4.302
- Sackett, P. R., Kuncel, N. R., Arneson, J. J., Cooper, S. R., & Waters, S. D. (2009). Does socioeconomic status explain the relationship between admissions tests and post-secondary academic performance? *Psychological Bulletin, 135*, 1–22. doi: 10.1037/a0013978
- Sackett, P. R., & Lievens, F. (2008). Personnel selection. *Annual Review of Psychology, 59*, 419–450. doi: 10.1146/annurev.psych.59.103006.093716
- Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. (2001). High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative-action world. *American Psychologist, 56*, 302–318. doi: 10.1037/0003-066X.56.4.302
- Salgado, J. F. (2002). The Big Five personality dimensions and counterproductive behaviors. *International Journal of Selection and Assessment, 10*, 117–125. doi: 10.1111/1468-2389.00198
- Salgado, J., & Anderson, N. (2002). Cognitive and GMA testing in the European community: Issues and evidence. *Human Performance, 15*, 75–96. doi: 10.1080/08959285.2002.9668084
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., & De Fruyt, F. (2003). International validity generalization of GMA and cognitive abilities: A European Community meta-analysis. *Personnel Psychology, 56*, 573–605. doi: 10.1111/j.1744-6570.2003.tb00751.x
- Schmitt, N. (2012). Development of rationale and measures of noncognitive college student potential. *Educational Psychologist, 47*, 18–29. doi: 10.1080/00461520.2011.610680
- Schmitt, N., Keeney, J., Oswald, F. L., Pleskac, T., Quinn, A., Sinha, R., & Zorrie, M. (2009). Prediction of 4-year college student performance using cognitive and noncognitive predictors and the impact of demographic status on admitted students. *Journal of Applied Psychology, 94*, 1479–1497. doi: 10.1037/a0016810
- Schmitt, N., Oswald, F. L., Friede, A., Imus, A., & Merritt, S. (2008). Perceived fit with an academic environment: Attitudinal and behavioral outcomes. *Journal of Vocational Behavior, 72*, 317–335. doi: 10.1016/j.jvb.2007.10.007

References

- Schmitt, N., Oswald, F. L., Kim, B. H., Gillespie, M. A., Ramsay, L. J., & Yoo, T. (2003). Impact of elaboration on social desirability and the validity of biodata measures. *Journal of Applied Psychology*, 88, 979–988. doi: 10.1037/0021-9010.88.6.979
- Schmitt, N., Rogers, W., Chan, D., Sheppard, L., & Jennings, D. (1997). Adverse impact and predictive efficiency of various predictor combinations. *Journal of Applied Psychology*, 82, 719–730. doi: 10.1037/0021-9010.82.5.719
- Schnabel, K. U., Alfeld, C., Eccles, J. S., Köller, O., & Baumert, J. (2002). Parental influence on students' educational choices in the United States and Germany: Different ramifications-same effect? *Journal of Vocational Behavior*, 60, 178–198. doi: 10.1006/jvbe.2001.1863
- Schutte, N. S., & Malouff, J. M. (2011). Emotional intelligence mediates the relationship between mindfulness and subjective well-being. *Personality and Individual Differences*, 50, 1116–1119. doi: 10.1016/j.paid.2011.01.037
- Schwager, I. T. L., Hülshager U. R., Bridgeman, B., & Lang, J. W. B. (2015). Graduate student selection: GRE, socioeconomic status and undergraduate grade point average as predictors of study success in a Western European university. *International Journal of Selection and Assessment*, 23, 71–79. doi: 10.1111/ijsa.12096
- Schwager, I. T. L., Hülshager U. R., Lang, J. W. B., Klieger, D. M., Bridgeman, B., & Wendler, C. (2014). Supervisor Ratings of Students' Academic Potential as Predictors of Citizenship and Counterproductive Behavior. *Learning and Individual Differences*, 35, 62–69. doi: 10.1016/j.lindif.2014.07.005
- Segal, Z. V., Williams, J. M. G., & Teasdale, J. D. (2002). *Mindfulness-based cognitive therapy for depression: A new approach to relapse prevention*. New York: Guilford Press.
- Shao, R., & Skarlicki, D. P. (2009). The role of mindfulness in predicting individual performance. *Canadian Journal of Behavioural Science*, 41, 195–201. doi: 10.1037/a0015166
- Shapiro, S. L., Brown, K. W., & Biegel, G. M. (2007). Teaching self-care to caregivers: effects of mindfulness-based stress reduction on the mental health of therapists in training. *Training and Education in Professional Psychology*, 1, 105–115. doi: 10.1037/1931-3918.1.2.105
- Shapiro, S. L., Schwartz, G. E., & Bonner, G. (1998). Effects of mindfulness-based stress reduction on medical and premedical students. *Journal of Behavioral Medicine*, 21, 581–599. doi: 10.1023/A:1018700829825
- Sherley, J. (2007). Letters: The utility of standardized tests. *Science*, 316, 1695–1696. Retrieved from <http://www.evsc.virginia.edu/wp-content/uploads/2007-LerdauAvery-Science.pdf>
- Sieverding, M., Schmidt, L. I., Obergfell, J., & Scheiter, F. (2013). Stress und Studienzufriedenheit bei Bachelor- und Diplomstudierenden im Vergleich: Eine Erklärung unter Anwendung des Demand-Control-Modells [Study-related stress and satisfaction among bachelor- and diploma-students: An explanation with the demand-control-model]. *Psychologische Rundschau*, 64, 94–100. doi: 10.1026/0033-3042/a000152

- Sijtsma, K. (2009). Reliability beyond theory and into practice. *Psychometrika*, 74, 169–173. doi: 10.1007/s11336-008-9103-y
- Sinha, R., Oswald, F., Imus, A., & Schmitt, N. (2011). Criterion-focused approach to reducing adverse impact in college admissions. *Applied Measurement in Education*, 24, 137–161. doi: 10.1080/08957347.2011.554605
- Smith, C. A., Organ, D. W., & Near, J. P. (1983). Organizational citizenship behavior: Its nature and antecedents. *Journal of Applied Psychology*, 68, 653–663. doi: 10.1037/0021-9010.68.4.653
- Spector, P. E., Bauer, J. A., & Fox, S. (2010). Measurement artifacts in the assessment of counterproductive work behavior and organizational citizenship behavior: Do we know what we think we know?. *Journal of Applied Psychology*, 95, 781–790. doi: 10.1037/a0019477.
- Stanton, B. C., Burstein, A. G., Kobos, J. C., & Loucks, S. (1979). The dean's letter of recommendation and resident performance. *Academic Medicine*, 54, 812–813. doi: 10.1097/00001888-197910000-00009
- Statistics Netherlands (2010). Many German students in Dutch higher education. Web magazine. Retrieved from <http://www.cbs.nl/nl-NL/menu/themas/onderwijs/publicaties/artikelen/archief/2010/2010-3022-wm.html>
- Staufenbiel, T., & Hartz, C. (2000). Organizational citizenship behavior: Development and validation of a measurement instrument. *Diagnostica*, 46, 73–83. doi: 10.1026//0012-1924.46.2.73
- Terry, L. S. 2008. The Bologna process and its impact in Europe: It's so much more than degree changes. *Vanderbilt Journal of Transnational Law*, 41, 107–228. Retrieved from <http://www.vanderbilt.edu/jotl/manage/wp-content/uploads/Terry-after-author-revisions-correct-final.pdf>
- Thompson, B. L., & Waltz, J. (2007). Everyday mindfulness and mindfulness meditation: Overlapping constructs or not?. *Personality and Individual Differences*, 43, 1875–1885. doi: 10.1016/j.paid.2007.06.017
- Trapmann, S., Hell, B., Hirn, J. O. W., & Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. *Zeitschrift für Psychologie/ Journal of Psychology*, 215, 132–151. Retrieved from <http://kops.uni-konstanz.de/handle/123456789/11149>
- Van de Werfhorst, H. G. (2009). Credential inflation and educational strategies: A comparison of the United States and the Netherlands. *Research in Social Stratification and Mobility*, 27, 269–284. doi: 10.1016/j.rssm.2009.10.001
- Verduin, J. R., & Clark, T. A. (1991). *Distance Education: The Foundations of Effective Practice*, Jossey-Bass Publishers, San Francisco, CA.
- Viswesvaran, C., & Ones, D. S. (2000). Perspectives on models of job performance. *International Journal of Selection and Assessment*, 8, 216–226. doi: 10.1111/1468-2389.00151
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81, 557–574. doi: 10.1037/0021-9010.81.5.557
- Walpole, M. B., Burton, N. W., Kanyi, K., & Jackenthal, A. (2001). *Selecting successful graduate students in-depth Interviews with GRE users*. Princeton, NJ: Educational Testing Service. Retrieved from <http://www.ets.org/Media/Research/pdf/RR-02-08-Walpole.pdf>

References

- Walters, A. M., Kyllonen, P. C., & Plante, J. W. (2006). Developing a standardized letter of recommendation. *Journal of College Admission*, 191, 8–17. Retrieved from <http://files.eric.ed.gov/fulltext/EJ741520.pdf>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, 54, 1063–1070. doi: 10.1037/0022-3514.54.6.1063
- Weinstein, N., Brown, K. W., & Ryan, R. M. (2009). A multi-method examination of the effects of mindfulness on stress attribution, coping, and emotional well-being. *Journal of Research in Personality*, 43, 374–385. doi: 10.1016/j.jrp.2008.12.008
- Weiss, H. M., & Cropanzano, R. (1996). Affective events theory: A theoretical discussion of the structure, causes and consequences of affective experiences at work. In B. M. Staw & L. L. Cummings (Eds.), *Research in organizational behavior: An annual series of analytical essays and critical reviews* (Vol. 18, pp. 1–74). Greenwich, CT: JAI Press.
- Wernimont, P. F., & Campbell, J. P. (1968). Signs, samples, and criteria. *Journal of Applied Psychology*, 52, 372–376. doi: 10.1037/h0026244
- Wolever, R. Q., Bobinet, K. J., McCabe, K., Mackenzie, E. R., Fekete, E., Kusnick, C. A., & Baime, M. (2012). Effective and viable mind-body stress reduction in the workplace: a randomized controlled trial. *Journal of Occupational Health Psychology*, 17, 246–258. doi: 10.1037/a0027278
- Woodhouse, B., & Jackson, P. H. (1977). Lower bounds for the reliability of the total score on a test composed of non-homogeneous items: II: A search procedure to locate the greatest lower bound. *Psychometrika*, 42, 579–591. doi: 10.1007/BF02295980
- Yorke, M. (2006). Employability in higher education: What it is what it is not. In M. Yorke (Ed.), *Learning and employability, Series 1*. New York: Learning and Teaching Support Network
- Zapf, D., Dormann, C., & Frese, M. (1996). Longitudinal studies in organizational stress research: A review of the literature with reference to methodological issues. *Journal of Occupational Health Psychology*, 1, 145–169. doi: 10.1037/1076-8998.1.2.145
- Zettler, I. (2011). Self-control and academic performance: Two field studies on university citizenship behavior and counterproductive academic behavior. *Learning and Individual Differences*, 21, 119–123. doi: 10.1016/j.lindif.2010.11.002
- Zwick, R. (2004). Is the SAT a “wealth test?” The link between educational achievement and socioeconomic status. In R. Zwick (Ed.) *Rethinking the SAT: The Future of Standardized Testing in University Admissions* (pp. 203–216). New York: RoutledgeFalmer.

Appendix

APPENDIX 1

Hierarchical Multiple Regression Analyses Predicting G-GPA from SES, UGPA, and GRE
Results for a Subsample with a Dutch Bachelor Degree

Predictor	G-GPA	
	ΔR^2	β
Step 1	.01	
SES		.11
Step 2	.14***	
U-GPA		.38***
Step 3	.03	
GRE-V		.08
GRE-Q		-.08
GRE-AW		.16
Total R^2	.18	

Note. $n = 136$. G-GPA = Graduate grade point average; SES = Mean score of father's and mother's education; U-GPA = Undergraduate grade point average; GRE-V = GRE Verbal Reasoning; GRE-Q = GRE Quantitative Reasoning; GRE-AW = GRE Analytical Writing.

*** $p < .001$

APPENDIX 2

Logistic Regression Analysis Predicting Degree Attainment within Scheduled Time Frame from SES, U-GPA, and GRE for a Subsample with a Dutch Bachelor Degree

Predictor	Model 1	Model 2	Model 3
SES	-.16		
U-GPA		.05	
GRE-V			.01
GRE-Q			.00
GRE-AW			-.30

Model fit			
χ^2	1.31	1.33	2.29
<i>df</i>	1	2	5
Nagelkerke R^2	0.01	0.01	0.02

Note. $n = 170$. SES = Mean score of father's and mother's education; U-GPA = Undergraduate grade point average; GRE-V = GRE Verbal Reasoning; GRE-Q = GRE Quantitative Reasoning; GRE-AW = GRE Analytical Writing. Degree attainment within scheduled time frame was coded 1 = finished in time; 0 = finished later than scheduled time frame.

APPENDIX 3

Hierarchical Multiple Regression Analyses Predicting University Citizenship and Counterproductive Academic Behavior from U-GPA and PPI Results for a Subsample with a Dutch Bachelor Degree

Predictor	University citizenship behavior	
	ΔR^2	β
Step 1	.01	
U-GPA		.09
Step 2	.14**	
Overall		.38**
Total R^2	.15**	
Counterproductive academic behavior		
Step 1	.05	
U-PGA		-.23
Step 2	.11**	
Overall		-.34**
Total R^2	.16**	

Note. $n = 81$. G-GPA = Graduate grade point average; SES = Mean score of father's and mother's education; U-GPA = Undergraduate grade point average; GRE-V = GRE Verbal Reasoning; GRE-Q = GRE Quantitative Reasoning; GRE-AW = GRE Analytical Writing.

** $p < .01$

Valorization

As part of the creation of the European Higher Education Area (EHEA) stated in the Bologna declaration, the objective has been formulated to promote students' long-term employability. This encompasses that students do not solely obtain the content knowledge related to their study field, but also obtain the skills required in a constantly changing world of work. The findings of this dissertation "It's not all about grades: New perspectives on graduate students' academic performance" provide meaningful insights in how to gain a student body that is both academically successful and valuable for the well-functioning of universities and future employers. Whereas Chapter 2 and 3 provide information about the predictive validity of a standardized cognitive test (Graduate Record Examination; GRE) and a noncognitive standardized letter of recommendation (Personal Potential Index; PPI), Chapter 4 encompasses initial empirical findings about the positive outcomes of a personality characteristic (trait mindfulness) having the potential to be trained. This means that universities could think about integrating mindfulness trainings into students' curriculum in order to improve their performance after they have been admitted to their study program. However, before the findings reported in Chapter 4 can be implemented into practice, further research is needed. Thus, in the remainder the focus will be on the two selection tools presented in Chapter 2 and 3 of this dissertation.

Given that ETS (Educational Testing Service) owns the copyright for the GRE and the PPI there is no direct possibility to profit monetarily from their implementation into practice. However, various stakeholders profit indirectly from selection tools covering a broader conceptualization of academic performance. Political stakeholders involved in the formulation of the Bologna protocol have emphasized the importance of noncognitive skills in order to improve students' value for the economy and the society. Likewise, universities in their mission statements stress the importance of both intellectual and personal development and growth. At the same time, students get the opportunity to receive a more complete picture of their cognitive abilities and noncognitive academic skills and can compensate for deficits on one or the other when applying for a master program. Thus, staff members involved in admission decisions at universities are in need of valid cognitive and noncognitive selection instruments that are fair, cost- and time-efficient.

Whereas the usefulness of cognitive admission tests such as the GRE is well-established in North America, findings presented in this dissertation are the first to confirm the usefulness of the GRE in the European context. Results reveal that students with higher GRE scores get better grades in their master program. The use of standardized tests is convenient from universities' perspective since test scores can be compared across applicants from various backgrounds. Especially due to students' increased mobility within Europe this is an important advantage over students' undergraduate grade point average (U-GPA) which is also widely used as an indicator of future study success. In order to predict students' noncognitive academic potential, the PPI provides meaningful information. Former supervisors are asked to evaluate the applicant on six prescribed dimensions such as teamwork performance or ethics and integrity. A written report is provided including the individual assessment per supervisor and an overall mean rating per dimension. Results of this dissertation reveal that students with high PPI evaluations are indeed more likely to engage in favorable behavior such as supporting fellow students and less likely to engage in detrimental behavior such as cheating or bullying.

The PPI as an alternative to traditional letters of recommendation is advantageous from universities' perspective as applicants' PPI evaluations can be easily rank ordered. Furthermore, an evaluation of the same dimensions is available across all applicants. In case study programs require specific noncognitive performance aspects (e.g., working in a team), universities/faculties can put different weight on the different PPI dimensions when taking admission decisions. From staff members' perspective the PPI is advantageous as filling in the PPI is far less time consuming than writing traditional letters of recommendation.

So far the different faculties at Maastricht University do not follow a consistent approach how to select their future master students. Most faculties select their master students based on their former grades, their content knowledge and their language proficiency (English and/or Dutch). Only some programs require standardized test results or ask students to write a motivation letter. However, a focus is clearly set on the assessment of students' cognitive potential. This is somehow surprising as the teaching system applied at Maastricht University (Problem Based Learning; PBL) explicitly requires students to function well in a group setting. During the tutorial group meetings students need to cooperate in order to cover all required study materials. Furthermore, they often have to work in small teams in order to prepare presentations, case studies or small projects. Thus, students showing cooperative instead counterproductive behavior are inevitable for the well-functioning of the PBL system. Students complaining about the PBL system often raise the point of critique that other students are free riding without being disciplined. This means that always the same students do the share of work whereas others are rather passive. By selecting students based on their noncognitive skills this problem could be reduced.

However, the assessment of cognitive skills should not be ignored. Especially in the Dutch higher education system the financial system is based on 'output funding' (that is: universities receive money for those students that have successfully graduated). This means that students who are not able to finish their master program cost a lot of money. Assuming that less skilled students require a substantial share of staff members' time and resources (e.g., they have to repeat an exam several times or need more supervision with their final thesis) this is especially problematic. Although being not skilled enough is not the only reason why students do not finish their study program, it is at least the one that can best be predicted beforehand. Thus, a wise decision should be taken whether quality or quantity is more important when admitting students to a master program. Taken together, we strongly encourage stakeholders involved in admission decisions at Maastricht University to think about a comprehensive selection system asking students to provide information about both their cognitive and noncognitive skills.

Summary

In response to the Bologna declaration the European higher education system has been comprehensively restructured within the last decade. One objective was to enhance students' employability to strengthen the competitiveness of the European labor market. Thus, especially on the graduate level where students are about to enter the job market, universities are encouraged to do both—transfer knowledge and promote students' noncognitive skills – in order to make them valuable job candidates. This is in line with the idea that academic as well as job performance encompasses three sub dimensions namely task performance, and citizenship and counterproductive behavior. Whereas task performance is best characterized by students' grades, students show citizenship behavior when they support their fellow students or engage in voluntary activities. Counterproductive behavior in contrast is negatively seen as it encompasses behaviors such as cheating or bullying. Thus, in order to obtain a student body with desirable outcomes on all three performance aspects, cognitive as well as noncognitive selection and assessment tools need to be considered.

Results presented in Chapter 2 of this dissertation reveal that the GRE, a standardized aptitude test developed to select students on the graduate level, predicts study success in terms of grades. Thus, before universities in the European context try to develop standardized tests themselves they should think about the advantages associated with the use of the GRE. Given that it is a well-established test for which predictive validity evidence already exists, universities do not risk being vulnerable to legal claims. Test scores can be easily compared and all students have equal chances to get permission to a master program. Depending on the specific content of the master program, universities might think about putting different weight on the three GRE parts.

Universities are not only interested in giving permission to those students that are most likely to obtain the highest grades. Students' noncognitive skills such as the ability to work in a group or to communicate determine both universities' well-functioning and students' sustainable employability. Of course it would be desirable to get an impression of applicants' noncognitive qualities by incorporating face-to-face interviews into the admission process. However, the associated costs are immense explaining the need to rely on alternative predictors of nontask performance. Results presented in Chapter 3 of this dissertation reveal that the PPI is a promising instrument predictive of citizenship and counterproductive behavior in the academic context. As former supervisors provide standardized information about students' noncognitive skills, the PPI overcomes a severe shortcoming associated with other noncognitive selection instruments namely social desirable responding.

Assuming that especially counterproductive academic behavior has far-reaching consequences while its prediction is not without difficulties, I successfully established the negative relationship between trait mindfulness and counterproductive academic behavior in Chapter 4 of this dissertation. Although students show dispositional differences with respect to their ability to nonjudgmentally experience the present moment, trait mindfulness has the advantage that it can be increased through training. Thus, universities might consider providing mindfulness trainings in order to reduce the likelihood that students engage in counterproductive behavior. However, results reveal that the positive effect of mindfulness remains within the boundaries of stable personality characteristics in such a way that especially students scoring low on conscientiousness and honesty-humility might profit from increased mindfulness levels.

Samenvatting

Het afgelopen decennium is het Europese onderwijssysteem erg veranderd als reactie op de Bologna verklaring. Eén van de doelen ervan was om de employability van studenten te vergroten om de competitie op de Europese arbeidsmarkt te versterken. Universiteiten worden om die reden aangemoedigd om zich bezig te houden met zowel kennisoverdracht, als ook het promoten van de niet-cognitieve vaardigheden van studenten, om ervoor te zorgen dat studenten waardevolle sollicitanten zullen zijn. Dit is met name van toepassing op het master niveau, waarbij studenten op korte termijn de arbeidsmarkt zullen betreden. Dit komt overeen met het idee dat zowel academische prestatie als werkprestatie drie verschillende sub dimensies bevatten, namelijk taak prestatie, citizenship behavior en contraproductief gedrag. Taak prestatie wordt gekenmerkt door de cijfers van studenten, terwijl studenten citizenship behavior vertonen in situaties waar zij hun medestudenten ondersteunen of deelnemen aan vrijwillige activiteiten. Contraproductief gedrag wordt echter gezien als negatief en het bevat gedragingen zoals afkijken of pesten. Om dus een studentenpopulatie te verkrijgen met gewenste uitkomsten op alle drie de aspecten, zal het nodig zijn om zowel cognitieve als niet-cognitieve selectie en assessment tools te overwegen.

De resultaten gepresenteerd in hoofdstuk 2 laten zien dat de GRE (Graduate Record Examination) - een gestandaardiseerde toets ontwikkeld om master studenten te selecteren - studie succes in de Europese context voorspelt. Dus, voordat universiteiten gaan proberen om zelf gestandaardiseerde testen te ontwikkelen zouden ze kunnen nadenken over de voordelen die verbonden zijn aan het gebruik van de GRE. Gezien het feit dat bewijsten aanzien van de predictieve validiteit al bestaat, lopen universiteiten niet het risico op juridische klachten. Testresultaten kunnen eenvoudig worden vergeleken en alle studenten hebben gelijke kansen om toelating voor een Masteropleiding te krijgen. Afhankelijk van de specifieke inhoud van een Masteropleiding, kunnen universiteiten overwegen om meer of minder nadruk te leggen op één van de drie GRE onderdelen.

Universiteiten zijn niet alleen geïnteresseerd om juist die studenten toe te laten waarvan de kans het grootst is dat ze hoge studiepunten zullen verkrijgen. Niet-cognitieve vaardigheden, zoals het vermogen om in een groep te functioneren of communicatieve vaardigheden, zijn ook belangrijk en hebben een grote invloed op het studiesucces van studenten. Natuurlijk zou het wenselijk zijn om gestandaardiseerde interviews in het toelatingsproces te integreren om een indruk van de sociale competenties van studenten te krijgen. Echter, de daaraan verbonden kosten zijn immens, waardoor universiteiten op zoek zijn naar alternatieve voorspellers. Hoofdstuk 3 geeft aan dat de PPI (Personal Potential Index) een veelbelovend instrument is om sociaal gedrag in de academische context te voorspellen. Aangezien supervisors de PPI beoordeling geven en niet de student zelf, is het probleem van sociaal wenselijke antwoorden opgelost.

Selectie is niet het enige geschikte middel om het goede functioneren van de studenten te optimaliseren. Ervan uitgaand dat de gevolgen van contraproductief academisch gedrag zoals afkijken, plagiaat of drugsmisbruik bijzonder ernstig zijn, is het aangetoonde verband met mindfulness dat gerapporteerd wordt in hoofdstuk 4 veelbelovend. Studenten met een hoog dispositioneel mindfulness niveau hebben minder kans om contraproductief gedrag laten zien. Gezien het feit dat het mindfulness niveau kan worden verbeterd door middel van training, worden universiteiten aanbevolen om mindfulness interventies in het studentenleven te integreren.

Acknowledgments

Liebe Ute, dir gilt in Bezug auf die letzten vier Jahre und die Erstellung meiner Doktorarbeit mein größter Dank. Ohne dich wäre das Ganze für mich undenkbar gewesen. Ich habe unfassbar viel von dir gelernt. Deine Art mich zu begleiten war für mich die perfekte Mischung aus Nachdruck und Vertrauen. Ich schätze dich sehr als Person in deinen vielfältigen Rollen als Wissenschaftlerin, (Doktor-) Mutter und Freundin. Ich werde die Zusammenarbeit mit dir sehr vermissen und wünsche mir sehr, dass wir auch in Zukunft in Kontakt bleiben werden.

Jonas ich danke dir, dass du es mir ermöglicht hast im Rahmen des ETS Projekts zu promovieren. Über weite Strecken waren wir Drei zweifelsohne ein gutes Team.

Dear Fred, thank you for always giving me the feeling that you will be there in case of an urgent matter.

Dear Brent and Cathy, thank you for giving me the opportunity to spend the time at ETS in Princeton. It was an unforgettable experience for me.

Liebe Katharina, ich bin ausgesprochen dankbar dich als Zimmerkollegin gehabt zu haben. Ich finde, dass wir eine sehr harmonische und schöne Zeit zusammen hatten.

Liebe Birthe, liebe Annika. Den Sommer 2014 werde ich immer mit euch und unseren Partybus-Fahrten zwischen Köln und Maastricht in Verbindung bringen.

Beste mensen (Alicia en Philippe), jullie twee zijn van begin aan in Maastricht daarbij geweest. Het was altijd heel vrolijk en gezellig met jullie.

Beste Mariella, het was altijd leuk om met jou over onze aanstaande vakantie plannen te kletsen.

Lieber Daniel, ohne deinen engagierten Einsatz als mein Bachelorarbeitsbetreuer in Aachen, hätte ich vermutlich niemals erst in Erwägung gezogen zu promovieren. Danke dafür!!!

Liebe Mama, lieber Papa, danke dass ihr immer für mich da seid und mir das Gefühl von Geborgenheit und Freiheit gebt. Ihr seid mein doppelter Boden und das werdet ihr auch für immer bleiben.

Liebe Andra, lieber Jan, ich liebe euch sehr und das nicht nur, weil das so zu sein hat, weil ihr meine Geschwister seid, sondern weil ich finde, dass ihr großartige Menschen seid. Danke dass ihr (und natürlich auch Ingo, Mathapelo und Vanessa) so wunderbare Kinder habt, die unser Familienleben unglaublich bereichern. Es tut sehr gut zu wissen, dass es euch allesamt gibt.

Liebste Patentante, liebste Lotti, die Yoga Ausflüge mit euch im letzten Jahr waren einfach nur wunderschön und haben mir mehr Energie gegeben als so mancher Jahresurlaub.

Liebste Katha, liebste Chantal – Mädels!!! Was wären die letzte vier Jahre nur ohne euch gewesen. Ich bin sehr froh, mit euch so verlässliche, fröhliche und tolle Frauen als beste Freundinnen an meiner Seite zu haben. Ich bin gespannt was die nächsten Jahre für uns bereithalten und freue mich sehr über die Gewissheit „dass ihr bleibt“.

Liebe Charly, schön dass es dich gibt und wir so eine witzige, abwechslungsreiche und wohlthuende Zeit zusammen haben.

Liebe Vanessa, du bist eine wirklich enge Freundin für mich geworden, deren ehrliche Meinung ich sehr zu schätzen weiß. Danke dafür!

Liebe Philippie, danke dir für deine Nachricht vom 04.Mai 2015.

Hvala Florijan Zivaljic! Die Endphase der Doktorarbeit werde ich immer mit dir in Verbindung bringen. Es war eine schöne Zeit mit dir und deiner Family.

Lieber Nico, auch wenn unsere gemeinsame Zeit in Aachen jetzt schon ein bisschen zurück liegt, gehörst du hier auch irgendwie hin. Gerade in der orientierungslosen Anfangsphase der Diss (so ungefähr die ersten drei Jahre ;-)) hast du mich sehr unterstützt. Dafür möchte ich dir danken.

Lieber Rolf, ich bin unglaublich froh, dass ich bei dir so ein wunderschönes Plätzchen in Köln gefunden habe. Ich fühle mich sehr zu Hause in der Clarenbachstrasse!

Curriculum Vitae

Inge Schwager was born on the 25th of November 1985 in Bad Honnef, Germany. After she obtained her Abitur at the Siebenbergsgymnasium Bad Honnef in 2005, she started studying Psychology at the Rheinische Friedrich-Wilhelms-University Bonn. In October 2007, she completed her Vordiplom and changed the university – after one year of studying at the RWTH Aachen, she obtained her Bachelor of Science in Psychology. In 2009 Inge started a master's degree program in Work- and Organizational Psychology at Maastricht University. After graduating as a Master of Science, she first worked as a Research Assistant in the Department of Work- and Social Psychology at Maastricht University followed by four year PhD project under the supervision of Dr. Ute Hülshager and Dr. Jonas Lang. The results of this project are presented in this dissertation. Inge starts working at the Deutsche Gesellschaft für Personalwesen (dgp e.V.) from the beginning of June 2015.